

```
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
```

```
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2
```

Query Match 85.9%, Score 1941.5, DB 4, Length 640;
Best Local Similarity 86.2%, Pred. No. 8.8e-145;
Matches 374, Conservative 21, Mismatches 38, Indels 1, Gaps 1;

QY 1 NDVARGIVKADVAONNGVLVGQGVAVVADTGLDTGRNDSMHEAFPGKITALVALGRTN 60
DB 207 NDVARGIVKADVAONNGVLVGQGVAVVADTGLDTGRNDSMHEAFPGKITALVALGRTN 266

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.2392 Seconds

(Without alignments) 2595.843 Million cell updates/sec

Title: US-09-985-689A-2

Perfect score: 2245

Sequence: 1 NDVARGIVKADVAQSSYGLY.....EYQAYNPVGPQNSLAIYN 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq19806:***
2: Geneseq19908:***
3: Geneseq20008:***
4: Geneseq20018:***
5: Geneseq20028:***
6: Geneseq20038:***
7: Geneseq20039:***
8: Geneseq20048:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2245	100.0	434	5 AAM50081	Aam50081 Bacillus
2	2237	99.6	639	2 AAY17089	Aay17089 Bacillus
3	2197	97.9	640	2 AAY17091	Aay17091 Bacillus
4	2181	97.6	434	5 AAM50080	Aam50080 Bacillus
5	2181	97.6	640	2 AAY17090	Aay17090 Bacillus
6	2155	96.0	639	2 AAY17087	Aay17087 An alkali
7	2155	96.0	640	2 AAY17088	Aay17088 An alkali
8	2130	94.9	434	5 AAM50085	Aam50085 Bacillus
9	2118.5	94.4	433	5 AAM50086	Aam50086 Bacillus
10	2118.5	94.4	641	2 AAM99547	Aam99547 Bacillus
11	2026	90.2	434	5 AAM50080	Aam50080 Bacillus
12	2010.5	89.6	433	5 AAM50084	Aam50084 Bacillus
13	2006.5	89.4	433	5 AAM50082	Aam50082 Bacillus
14	2004.5	89.3	636	2 AAM89548	Aam89548 Bacillus
15	1999.5	89.1	433	5 AAM50083	Aam50083 Bacillus
16	1998.5	89.0	433	5 AAR6274	Aar6274 Alkali-pr
17	1998.5	89.0	433	2 AAM61495	Aam61495 Modified
18	1998.5	89.0	433	2 AAM56588	Aam56588 Bacillus
19	1998.5	89.0	433	3 AAY69207	Aay69207 Amino aci
20	1998.5	89.0	433	3 AAY44619	Aay44619 Bacillus
21	1579.5	70.4	345	2 AAM62230	Aam62230 Subtilase
22	1579.5	70.4	345	2 AAY21654	Aay21654 Subtilase
23	457.5	20.4	659	2 AAM24121	Aam24121 Thermococ
24	457.5	20.4	659	2 AAM94840	Aam94840 MO9856926
25	423	18.8	412	2 AAM94836	Aam94836 Hyperther

26	423	18.8	522	2 AAM24122	Aam24122 Pyrococcus
27	423	18.8	522	2 AAM94838	Aam94838 Hyperther
28	423	18.8	654	2 AAM24129	Aam24129 Pyrococcus
29	423	18.8	654	2 AAM94841	Aam94841 Hyperther
30	412	18.4	659	2 AAM24123	Aam24123 Protease.
31	383	17.1	545	4 AAB09483	Abb09483 T. yonsei
32	365	16.3	1079	6 AAB81180	Abb81180 Transglut
33	365	16.3	1079	6 AABU07391	Abu07391 Foreign p
34	354.5	15.8	520	2 AAM13666	Aam13666 Fragment
35	354.5	15.8	734	2 AAM13667	Aam13667 Streptom
36	354.5	15.8	823	2 AAM13668	Aam13668 DHPA-mel
37	343	15.3	1237	6 AABU11343	Abu11343 Protein e
38	316.5	14.1	699	2 AAY08471	Aay08471 F. balust
39	308.5	13.7	519	6 AABP76735	Abp76735 Streptom
40	308.5	13.7	19938	6 AABP76678	Abp76678 Streptom
41	304.5	13.6	903	2 AAR87007	Aar87007 Hyperther
42	304.5	13.6	1398	2 AAR87008	Aar87008 Protease.
43	304.5	13.6	1398	2 AAM24124	Aam24124 Pyrococcus
44	304.5	13.6	1398	2 AAM94839	Aam94839 MO9856926
45	302.5	13.5	806	2 AAR27481	Aar27481 RP-III re

ALIGNMENTS

RESULT 1	AAM50081	standard; protein; 434 AA.
ID	AAM50081	
XX	AC	AAM50081;
XX	DT	12-AUG-2002 (first entry)
XX	DE	Bacillus sp KSM-KP9860 alkaline protease protein fragment.
XX	KW	Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX	OS	Bacillus sp.
XX	PN	EP1209233-A2.
XX	PD	29-MAY-2002.
XX	PF	22-NOV-2001; 2001EP-00127851.
XX	PR	22-NOV-2000; 2000JP-00355166.
XX	PR	12-APR-2001; 2001JP-00114048.
XX	PA	(KAOS) KAO CORP.
XX	PI	Harada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX	FI	Okuda M, Saeki K;
XX	DR	WPI; 2002-437518/47.
XX	PT	New modified alkaline proteases useful in detergent compositions.
XX	PS	Claim 5; Page 12-13; 25pp; English.
XX	CC	This invention describes novel Bacillus sp. alkaline proteases useful in
XX	CC	detergent compositions, especially in laundry, bleaching or automatic
XX	CC	dishwasher detergents. The novel proteases have an increased detergency &
XX	CC	(34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX	CC	sequence represents a fragment of the alkaline protease KP9860 from
XX	CC	Bacillus sp strain KSM-KP9860 described in the method of the invention
XX	SO	Sequence 434 AA;
XX	QY	Query Match 100.0%; Score 2245; DB 5; Length 434;
XX	QY	Best Local Similarity 100.0%; Pred. No. 1e-155;
XX	QY	Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	QY	1 NDVARGIVKADVAQSSYGLYGGQGIYAVADTGTDTGRNDSMHEARFKITAIYALGRIN 60

```

Db      1 NDVARGIVKADVAQSSVGLYGQGIYAVADTGIDTGRNDSMHEARFGKITLALGRN 60
Qy      61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFOSIMDSGGIGLPSNLQTLFSGAFS 120
Db      61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFOSIMDSGGIGLPSNLQTLFSGAFS 120
Qy      121 AGARHNTSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEPNGGTTISAPGTAKNAI 180
Db      121 AGARHNTSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEPNGGTTISAPGTAKNAI 180
Qy      181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAEGTYILSRSSLPDSSF 240
Db      181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAEGTYILSRSSLPDSSF 240
Qy      241 WANHDSKYAVMGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALIIAGADVGLGY 300
Db      241 WANHDSKYAVMGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALIIAGADVGLGY 300
Qy      301 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFETATAGKPLKISLVMSDAPASTTA 360
Db      301 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFETATAGKPLKISLVMSDAPASTTA 360
Qy      361 SVTLVNDLDELVTAPNGTRVYGNDFSAFPDNNMDGRNVENVFINSPOSGTYTIEVOAYN 420
Db      361 SVTLVNDLDELVTAPNGTRVYGNDFSAFPDNNMDGRNVENVFINSPOSGTYTIEVOAYN 420
Qy      421 VPVGPQNFSLAIYN 434
Db      421 VPVGPQNFSLAIYN 434

```

RESULT 2
AA17089
ID AA17089 standard; protein; 639 AA.

```

XX      AA17089;
AC      20-MAR-2003 (revised)
DI      21-JUL-1999 (first entry)
XX      DE Bacillus alkaline protease.
XX      KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX      washing composition; oxidizing agent.
XX      OS Bacillus sp.
XX      PN W09918218-A1.
XX      PD 15-APR-1999.
XX      PF 07-OCT-1998; 98WO-JP004528.
XX      PR 07-OCT-1997; 97JP-00274570.
XX      (KAOS ) KAO CORP.
XX      PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
XX      PI Shikata S, Nomura M,
XX      DR WPI; 1999-287736/27.
XX      DR N-PSDB; AAX37277.
XX      PT Alkali protease from Bacillus used in washing powders.
XX      PS Disclosure; Page 53-58; 71pp; Japanese.
XX      CC The invention relates to alkaline proteases produced by strains of
XX      CC Bacillus; the proteases ability to digest casein is not inhibited by
XX      CC oleic acid and they have a high stability to oxidizing agents. The
XX      CC alkaline protease of the invention has the following properties: (a) it
XX      CC is active over the pH range 4-13 and has at least 80% of its optimum

```

CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d) its ability to digest casein is not inhibited by oleic acid; (e) it has molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be used as enzymes in washing compositions for use in automatic dishwashers and for washing clothes. The stability to oxidizing agents allows the enzyme to be an effective component of washing compositions including bleach. The present sequence represents an alkaline protease. (updated on 20-MAR-2003 to correct DR field.)

XX SQ Sequence 639 AA;

Query Match 99.6%; Score 2237; DB 2; Length 639;
Best Local Similarity 99.8%; Pred. No. 6.6e-155;
Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 NDVARGIVKADVAQSSVGLYGQGIYAVADTGIDTGRNDSMHEARFGKITLALGRN 60
Db      206 NDVARGIVKADVAQSSVGLYGQGIYAVADTGIDTGRNDSMHEARFGKITLALGRN 265
Qy      61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFOSIMDSGGIGLPSNLQTLFSGAFS 120
Db      266 NANDTNGHGHVAGSVLGNATNKGAPOANLVFOSIMDSGGIGLPSNLQTLFSGAFS 325
Qy      121 AGARHNTSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEPNGGTTISAPGTAKNAI 180
Db      326 AGARHNTSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNERPNGGTTISAPGTAKNAI 385
Qy      181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAEGTYILSRSSLPDSSF 240
Db      386 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAEGTYILSRSSLPDSSF 445
Qy      241 WANHDSKYAVMGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALIIAGADVGLGY 300
Db      446 WANHDSKYAVMGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALIIAGADVGLGY 505
Qy      301 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFETATAGKPLKISLVMSDAPASTTA 360
Db      506 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFETATAGKPLKISLVMSDAPASTTA 565
Qy      361 SVTLVNDLDELVTAPNGTRVYGNDFSAFPDNNMDGRNVENVFINSPOSGTYTIEVOAYN 420
Db      566 SVTLVNDLDELVTAPNGTRVYGNDFSAFPDNNMDGRNVENVFINSPOSGTYTIEVOAYN 625
Qy      421 VPVGPQNFSLAIYN 434
Db      626 VPVGPQNFSLAIYN 639

```

RESULT 3
AA17091
ID AA17091 standard; protein; 640 AA.

```

XX      AA17091;
AC      20-MAR-2003 (revised)
DI      21-JUL-1999 (first entry)
XX      DE Bacillus alkaline protease.
XX      KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX      KM washing composition; oxidizing agent.
XX      OS Bacillus sp.
XX      PN W09918218-A1.
XX      PD 15-APR-1999.
XX      PF 07-OCT-1998; 98WO-JP004528.
XX      PR 07-OCT-1997; 97JP-00274570.
XX

```

PA (KAOS) KAO CORP.
 XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
 PI Shikata S, Nomura M;
 XX MPI: 1999-287726/27.
 DR N-PSDB; AAK37279.
 XX Alkali protease from *Bacillus* used in washing powders.
 PS Disclosure; Page 63-68; 71pp; Japanese.
 XX
 CC The invention relates to alkaline proteases produced by strains of
 CC *Bacillus*. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidizing agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 By SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidizing agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease. (Updated
 CC on 20-MAR-2003 to correct DR field.)
 CC
 SQ Sequence 640 AA:
 Query Match 97.9%; Score 2197; DB 2; Length 640;
 Best Local Similarity 96.8%; Pred. No. 5.6e-152;
 Matches 440; Conservative 13; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLALGRIN 60
 DB 207 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLALGRIN 266
 QY 61 NNDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
 DB 267 NNDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 326
 QY 121 AGARITNSGGAANGAYTTDSRNVDDYRKNDMTILFAAGNEGNGGTTISAPGTAKNAI 180
 DB 327 AGARITNSGGAANGAYTTDSRNVDDYRKNDMTILFAAGNEGNGGTTISAPGTAKNAI 386
 QY 181 TVGATENLRFSGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 240
 DB 387 TVGATENLRFSGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 446
 QY 241 WANHDSKAYVNGGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGAADVGLGY 300
 DB 447 WANHDSKAYVNGGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGAADVGLGY 506
 QY 301 PNGNGMGWVTLDKSLNVAIVNNESSLSISQKATYFTATAGKPKLISLWSDAPASTTA 360
 DB 507 PNGNGMGWVTLDKSLNVAIVNNESSLSISQKATYFTATAGKPKLISLWSDAPASTTA 566
 QY 361 SVTLVNDLDELVTAPNGTRYVGNDFAPDNMWDGNNVENVFINSPOGTYTIEVQAYN 420
 DB 567 SVTLVNDLDELVTAPNGTRYVGNDFAPDNMWDGNNVENVFINSPOGTYTIEVQAYN 626
 QY 421 VVVGPNQFSLATVN 434
 DB 627 VVVGPNQFSLATVN 640

RESULT 4
 AAM50080
 ID AAM50080 standard; protein; 434 AA.
 XX AAM50080;
 AC
 XX
 DT 12-AUG-2002 (first entry)

XX
 DE *Bacillus* sp KSM-KP43 alkaline protease protein fragment.
 XX
 KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS *Bacillus* sp.
 XX
 PN EPI209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 XX 22-NOV-2001; 2001EP-00127851.
 PF
 XX 22-NOV-2000; 2000JP-00355166.
 PR
 XX 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Arai H, Sumitomo N,
 PI Okuda M, Saeki K;
 XX MPI: 2002-437518/47.
 DR
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 1; Page 10-11; 25pp; English.
 XX
 CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 28%). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC *Bacillus* sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50090
 CC
 SQ Sequence 434 AA:
 Query Match 97.6%; Score 2191; DB 5; Length 434;
 Best Local Similarity 96.5%; Pred. No. 9.1e-152;
 Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLALGRIN 60
 DB 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLALGRIN 60
 QY 61 NNDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
 DB 61 NNDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
 QY 121 AGARITNSGGAANGAYTTDSRNVDDYRKNDMTILFAAGNEGNGGTTISAPGTAKNAI 180
 DB 121 AGARITNSGGAANGAYTTDSRNVDDYRKNDMTILFAAGNEGNGGTTISAPGTAKNAI 180
 QY 181 TVGATENLRFSGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 240
 DB 181 TVGATENLRFSGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSF 240
 QY 241 WANHDSKAYVNGGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGAADVGLGY 300
 DB 241 WANHDSKAYVNGGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGAADVGLGY 300
 QY 301 PNGNGMGWVTLDKSLNVAIVNNESSLSISQKATYFTATAGKPKLISLWSDAPASTTA 360
 DB 301 PNGNGMGWVTLDKSLNVAIVNNESSLSISQKATYFTATAGKPKLISLWSDAPASTTA 360
 QY 361 SVTLVNDLDELVTAPNGTRYVGNDFAPDNMWDGNNVENVFINSPOGTYTIEVQAYN 420
 DB 361 SVTLVNDLDELVTAPNGTRYVGNDFAPDNMWDGNNVENVFINSPOGTYTIEVQAYN 420
 QY 421 VVVGPNQFSLATVN 434
 DB 421 VVVGPNQFSLATVN 434

XX Sequence 639 AA;
SQ Query Match 96.0%; Score 2155; DB 2; Length 639;
Best Local Similarity 96.3%; Pred. No. 6,5e-149;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 265
QY 61 NANDTNGHGTIVAGSVLGNATNKGAPOANLVFQSIIMDSGGIGLPENLQTLFSGQAFS 120
DB 266 NANDTNGHGTIVAGSVLGNATNKGAPOANLVFQSIIMDSGGIGLPENLQTLFSGQAFS 325
QY 121 AGARHTNSMGAANVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 180
DB 326 AGARHTNSMGAANVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 385
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSSILAPDSSF 240
DB 386 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSSILAPDSSF 445
QY 241 WANHDSKRYAVMGSTSMATPIVAGNVAQLREHFYKNGITPKPSILKALILAGAADVGLGY 300
DB 446 WANHDSKRYAVMGSTSMATPIVAGNVAQLREHFYKNGITPKPSILKALILAGAADVGLGY 505
QY 301 PNGNQGMGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 360
DB 506 PNGNQGMGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 565
QY 361 SVTLVNDLDLVITAPNGTRVGNDFSAFPDMMDGNNVENVFINSQSGTITIEVOAYN 420
DB 566 SVTLVNDLDLVITAPNGTRVGNDFSAFPDMMDGNNVENVFINSQSGTITIEVOAYN 625
QY 421 VPVGPOXFSFLAIYN 434
DB 626 VPVGPOXFSFLAIYN 639
RESULT 7
ID AAY17088 standard; protein; 640 AA.
XX AAY17088;
AC 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE An alkaline protease sequence from Bacillus species.
XX
KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent.
XX
OS Bacillus sp.
XX
XX Key Location/Qualifiers
FT Misc-difference 1..640
FT /note= "all residues indicated as Xaa are arbitrary amino
FT acids"
XX
XX W09918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
PI Shikata S, Nomura M,
PI

XX WP1: 1999-287736/27.
DR N-PSDB; AAX37278.
XX
FT Alkali protease from Bacillus used in washing powders.
XX
PS Claim 3; Page 50-53; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 640 AA;
QY Query Match 96.0%; Score 2155; DB 2; Length 640;
Best Local Similarity 96.3%; Pred. No. 6,5e-149;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 207 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
QY 61 NANDTNGHGTIVAGSVLGNATNKGAPOANLVFQSIIMDSGGIGLPENLQTLFSGQAFS 120
DB 267 NANDTNGHGTIVAGSVLGNATNKGAPOANLVFQSIIMDSGGIGLPENLQTLFSGQAFS 326
QY 121 AGARHTNSMGAANVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 180
DB 327 AGARHTNSMGAANVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 386
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSSILAPDSSF 240
DB 387 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSSILAPDSSF 446
QY 241 WANHDSKRYAVMGSTSMATPIVAGNVAQLREHFYKNGITPKPSILKALILAGAADVGLGY 300
DB 447 WANHDSKRYAVMGSTSMATPIVAGNVAQLREHFYKNGITPKPSILKALILAGAADVGLGY 506
QY 301 PNGNQGMGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 360
DB 507 PNGNQGMGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 566
QY 361 SVTLVNDLDLVITAPNGTRVGNDFSAFPDMMDGNNVENVFINSQSGTITIEVOAYN 420
DB 567 SVTLVNDLDLVITAPNGTRVGNDFSAFPDMMDGNNVENVFINSQSGTITIEVOAYN 626
QY 421 VPVGPOXFSFLAIYN 434
DB 627 VPVGPOXFSFLAIYN 640
RESULT 8
ID AAM50085 standard; protein; 434 AA.
XX AAM50085;
AC 12-AUG-2002 (first entry)
DT
XX
XX Bacillus sp alkaline protease protein A-1 fragment.
DE
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
KW

XX OS Bacillus sp.
 XX PN EPI209233-A2.
 XX PD 29-MAY-2002.
 XX PF 22-NOV-2001; 2001EP-00127851.
 XX PR 22-NOV-2000; 2000JP-00355166.
 XX PR 12-APR-2001; 2001JP-00114048.
 XX PA (KAO) KAO CORP.
 XX PI Hatada Y, Ogawa A, Kagayama Y, Sato T, Araki H, Sumitomo N;
 XX PI Okuda M, Saeki K;
 XX DR WPI; 2002-437518/47.
 XX PT New modified alkaline proteases useful in detergent compositions.
 XX PS Claim 5; Page 18-19; 25pp; English.
 XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency *
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease A-1 from Bacillus
 CC sp NCIB12289 described in the method of the invention

XX SQ Sequence 434 AA;

Query Match 94.9%; Score 2130; DB 5; Length 434;
 Best Local Similarity 92.4%; Pred. No. 2.6e-147;
 Matches 401; Conservative 24; Mismatches 9; Indels 0; Gaps 0;

DB 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITATLYALGRIN 60
 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITATLYALGRIN 60
 QY 61 NANDTNHGHGTHVAGSVLGNATNKMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGAFS 120
 DB 61 NANDPNHGHGTHVAGSVLGNATNKMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGAFS 120
 QY 121 AGARIHNSWGAANVAGVYTTDSRVVDYVRKNMTILFAAGNEGPNGGTSAPGTAKNAI 180
 DB 121 AGARIHNSWGAANVAGVYTTDSRVVDYVRKNMTILFAAGNEGPNGGTSAPGTAKNAI 180
 QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIPDVMAPEQTYILSARSLAPDSGF 240
 DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIPDVMAPEQTYILSARSLAPDSGF 240
 QY 241 WANDSKAYWGGSMATPIVAGVNAQLREHFVNRGITPKPSILKALTAGADVGLG 300
 DB 241 WANDSKAYWGGSMATPIVAGVNAQLREHFVNRGITPKPSILKALTAGADVGLG 300
 QY 301 PNGNQGWGRVTLDSKLNVAVYVESALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
 DB 301 PNGNQGWGRVTLDSKLNVAVYVESALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
 QY 361 SVTLVNDLDDVITAPNGRKYGNDFSAFDDNMGRNNVENVFINSQSGTYTVEQAYN 420
 DB 361 SVTLVNDLDDVITAPNGRKYGNDFSAFDDNMGRNNVENVFINSQSGTYTVEQAYN 420
 QY 421 VPVGPNFSLAIYN 434
 DB 421 VPVGPNFSLAIYN 434

RESULT 9
 AAMS0086
 ID AAMS0086 standard; protein; 433 AA.
 XX

AC AAMS0086;
 XX 12-AUG-2002 (first entry)
 DT Bacillus sp alkaline protease protein A-2 fragment.
 XX DE Bacillus sp alkaline protease protein A-2 fragment.
 XX DE Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX KW Bacillus sp.
 XX OS EPI209233-A2.
 XX PN 29-MAY-2002.
 XX PD 22-NOV-2001; 2001EP-00127851.
 XX PF 22-NOV-2000; 2000JP-00355166.
 XX PR 12-APR-2001; 2001JP-00114048.
 XX PA (KAO) KAO CORP.
 XX PI Hatada Y, Ogawa A, Kagayama Y, Sato T, Araki H, Sumitomo N;
 XX PI Okuda M, Saeki K;
 XX DR WPI; 2002-437518/47.
 XX PT New modified alkaline proteases useful in detergent compositions.
 XX PS Claim 5; Page 20-21; 25pp; English.

CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency *
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease A-2 from Bacillus
 CC sp NCIB12513 described in the method of the invention

XX SQ Sequence 433 AA;

Query Match 94.4%; Score 2118.5; DB 5; Length 433;
 Best Local Similarity 93.3%; Pred. No. 1.8e-146;
 Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITATLYALGRIN 60
 DB 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITATLYALGRIN 60
 QY 61 NANDTNHGHGTHVAGSVLGNATNKMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGAFS 120
 DB 61 NANDPNHGHGTHVAGSVLGNATNKMAPOANLVFQSIMDSGGGLGGLPSNLQTLFSGAFS 119
 QY 121 AGARIHNSWGAANVAGVYTTDSRVVDYVRKNMTILFAAGNEGPNGGTSAPGTAKNAI 180
 DB 121 AGARIHNSWGAANVAGVYTTDSRVVDYVRKNMTILFAAGNEGPNGGTSAPGTAKNAI 179
 QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIPDVMAPEQTYILSARSLAPDSGF 240
 DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIPDVMAPEQTYILSARSLAPDSGF 239
 QY 241 WANDSKAYWGGSMATPIVAGVNAQLREHFVNRGITPKPSILKALTAGADVGLG 300
 DB 241 WANDSKAYWGGSMATPIVAGVNAQLREHFVNRGITPKPSILKALTAGADVGLG 299
 QY 301 PNGNQGWGRVTLDSKLNVAVYVESALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
 DB 301 PNGNQGWGRVTLDSKLNVAVYVESALSTSOKATYFTATAGKPLKISLWSDAPASTTA 359
 QY 361 SVTLVNDLDDVITAPNGRKYGNDFSAFDDNMGRNNVENVFINSQSGTYTVEQAYN 420
 DB 361 SVTLVNDLDDVITAPNGRKYGNDFSAFDDNMGRNNVENVFINSQSGTYTVEQAYN 419
 QY 421 VPVGPNFSLAIYN 434
 DB 421 VPVGPNFSLAIYN 434

DB 420 VPVSPQTFSLAIHV 433

RESULT 10

AAW89547 ID AAW89547 standard; protein; 641 AA.

XX AAW89547;

XX AC

XX DT 12-APR-1999 (first entry)

XX DE Bacillus JPI70 protease.

XX KM Protease; detergent; surfactant; leather processing; deblistering; flavour.

XX OS Bacillus sp.

XX Key Location/Qualifiers

FT Peptide 1..33

FT Region /note= "signal peptide"

FT /note= "prepro region"

FT Protein 209..641

FT /note= "mature protein"

XX W09856927-A2.

XX 17-DEC-1998.

XX PF 09-JUN-1998; 98WO-US012005.

XX PR 12-JUN-1997; 97US-00873479.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PI Sloma A, Christianson L;

XX WPI: 1999-080908/07.

XX N-PSDB; AAW82382.

XX Novel protease from Bacillus subtilis LC20 - useful in laundry and dishwashing detergents and for leather processing.

XX PS Claim 7, Page 53-54; 77pp; English.

XX This is the amino acid sequence of a novel protease of Bacillus sp. JPI70 (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene (see AAW82382). The entire protein, including the signal peptide and prepro region, has 77% identity to alkaline protease Y (see AAW89548) from Bacillus. The invention provides vectors, recombinant host cells and methods for the recombinant production of the protease. The protease is used in laundry and dishwashing detergents, for institutional and industrial cleaning, and for leather processing, as well as for deblistering and enhancing the degree of hydrolysis of proteins, hydrolyzates, for flavour development through hydrolysis of proteins, degradation of undesired peptides and in enzymatic synthesis of peptides. It has enhanced stability towards oxidation under alkaline conditions, e.g. towards bleaching agents of the peroxy type. The invention also provides mutant cells in which the protease activity is diminished. Such cells can be used for the production of heterologous recombinant proteins

XX Sequence 641 AA;

XX SQ

Query Match 94.4%; Score 2118.5; DB 2; Length 641;

Best Local Similarity 93.3%; Pred. No. 3e-146;

Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;

QY 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDGRNDSNHEAFRGKITLALYLGITN 60

DB 209 NDVARGIVADVAQNNFGLYGQGIYAVADTGLDGRNDSNHEAFRGKITLALYLGITN 268

QY 61 NNDNTGHTHVAGSVLGNATNKGAPANLVFQISIMDSGGGLGSLPNTQLTFSQAFS 120

DB 269 NNDPENGHTHVAGSVLGN-ATNKGAPANLVFQISIMDSGGGLGSLPNTQLTFSQAFS 327

QY 121 AGARITNSWGAANVAGYTTDSRNDVDYRKNDMTILFAAGNEGNGGTISAPGAKNAI 180

DB 328 AGARITNSWGAAPVNGAYTTDSRNDVDYRKNDMTILFAAGNEGNGGTISAPGAKNAI 387

QY 181 TVGATENLRPSFGSYADNINNHVAFSSRGPTKDGRIKPDVMAPGTYILSARSSLA 240

DB 388 TVGATENLRPSFGSYADNINNHVAFSSRGPTKDGRIKPDVMAPGTYILSARSSLA 447

QY 241 MANHDSKIAYMGTSNATPIVAGNVAQLREHVYKRGITPKSLKALIAAAYGLGY 300

DB 448 MANHDSKIAYMGTSNATPIVAGNVAQLREHVYKRGITPKSLKALIAAAYGLGY 507

QY 301 PNGNGMGKVTLDKSLNVAAYVNESSALSTGKATYFTATAGKPLKISLWMSDAPASTTA 360

DB 508 PNGNGMGKVTLDKSLNVAAYVNESSALSTGKATYFTATAGKPLKISLWMSDAPASTTA 567

QY 361 SVTLVNDLVLVITAPNGTRYVGNDSAPPDNNMDGRNNVENVFINSPOSGYTTIVQAYN 420

DB 568 SVTLVNDLVLVITAPNGTRYVGNDSAPPDNNMDGRNNVENVFINSPOSGYTTIVQAYN 627

QY 421 VPVGPQNFSLATVN 434

DB 628 VPVSPQTFSLAIHV 641

RESULT 11

AAW50090 ID AAW50090 standard; protein; 434 AA.

XX AAW50090;

XX 12-AUG-2002 (first entry)

XX DE Bacillus sp KSW-KP43 alkaline protease protein variant.

XX KM Alkaline protease; detergent; laundry; bleaching; dishwasher; mutant;

XX KM mutcin.

XX OS Bacillus sp.

XX OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 46

FT /label= 'y,w,a,d,e,t,v,l,i,h,s,k,q,m,c

FT /note= "as claimed in Claim 3"

FT MISC-difference 54

FT /label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c

FT /note= "as claimed in Claim 3"

FT MISC-difference 57

FT /label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a

FT /note= "as claimed in Claim 3"

FT MISC-difference 66

FT /label= 'e,d,s,g,a,t,i,m,c,v,g,i

FT /note= "as claimed in Claim 3"

FT MISC-difference 84

FT /label= OTHER, R

FT /note= "OTHER= deleted residue. Specifically described in Claim 1"

FT MISC-difference 101..106

FT /label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a

FT /note= "as claimed in Claim 3"

FT MISC-difference 104

FT /label= OTHER, P

FT /note= "OTHER= deleted residue. Specifically described in Claim 1"

FT MISC-difference 107

FT /label= 'k,r,a,b

FT /note= "as claimed in Claim 3"

FT MISC-difference 119

FT /label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c

FT /note="as claimed in Claim 3"
 FT Misc-difference 124 /label="a,k
 FT /note="as claimed in Claim 3"
 FT Misc-difference 136 /label="k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
 FT /note="as claimed in Claim 3"
 FT Misc-difference 138 /label="y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
 FT /note="as claimed in Claim 3"
 FT Misc-difference 148 /label="y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
 FT /note="as claimed in Claim 3"
 FT Misc-difference 193 /label="k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
 FT /note="as claimed in Claim 3"
 FT Misc-difference 195 /label="y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
 FT /note="as claimed in Claim 3"
 FT Misc-difference 205 /label="y,w,a,d,e,t,v,l,i,h,s,k,q,m,c
 FT /note="as claimed in Claim 3"
 FT Misc-difference 247 /label="w,f,a,r,e,t,v,l,i,h,s,e,m,c
 FT /note="as claimed in Claim 3"
 FT Misc-difference 256 /label="OTHER=a,s,e,v,l,r,e,d
 FT /note="OTHER= deleted residue. Specifically described in Claim 1"
 FT Misc-difference 257 /label="v,i
 FT /note="as claimed in Claim 3"
 FT Misc-difference 264 /label="e,d,s,g,a,t,l,m,c,v,g,i
 FT /note="as claimed in Claim 3"
 FT Misc-difference 342 /label="k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
 FT /note="as claimed in Claim 3"
 FT Misc-difference 369 /label="OTHER, d
 FT /note="OTHER= deleted residue. Specifically described in Claim 1"
 EP1209233-A2.
 XX 29-MAY-2002.
 XX 22-NOV-2001; 2001EP-00127851.
 XX 22-NOV-2000; 2000JP-00355166.
 XX 12-APR-2001; 2001JP-00114048.
 XX (KAOS) KAO CORP.
 XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX WPI; 2002-437518/47.
 DR New modified alkaline proteases useful in detergent compositions.
 PT Claim 1; Page; 25pp; English.
 XX This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a variant of the alkaline protease Kp43 from *Bacillus*
 CC sp strain KSM-KP43 created from the wild-type protease represented in
 CC AAM50080 NOTE: This sequence is not represented in the specification but
 CC has been constructed from the sequence represented in record AAM50080
 XX
 SO Sequence 434 AA;

Query Match 90.2%; Score 2026; DB 5; Length 434;
 Best Local Similarity 91.0%; Pred. No. 1e-139;
 Matches 395; Conservative 12; Mismatches 27; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQSSGYGLYGQGIYAVADTGDTGRNDSMHEARFGKITLALGRTN 60
 DB 1 NDVARGIVKADVAQSSGYGLYGQGIYAVADTGDTGRNDSMHEARFGKITLALGRTN 60
 QY 61 NADNTNGHGTIVAGSVLNGATNKGAPOANLVFQSIMDSGGELGGLPSNLQTLFQAFS 120
 DB 61 NADNTXGHGTIVAGSVLNGSTYKXGMAPQANLVFQSIMDSXXXXXXPSNLQTLFQAFS 120
 QY 121 AGARITNMSGAAVNGATYDTSRRVDYRKQNTLTFPAAGNCPGKTIAPGTXKAI 180
 DB 121 AGARITNMSGAAVNGATYDTSRRVDYRKQNTLTFPAAGNCPGKTIAPGTXKAI 180
 QY 181 TVGATENLRPSFGSYADNINHVAQFSRGPETKDGRIKPDVMAFGTYILSRSSLABDSF 240
 DB 181 TVGATENLRPSFGSXADNINHVAQFSRGPETKDGRIKPDVMAFGTYILSRSSLABDSF 240
 QY 241 WANHDSKYATMGTSMAPIVAAGVAQREHFYKRGITPKBSLLKALILAGADVGLG 300
 DB 241 WANHDSKYATMGTSXTPVAAGVAQREHFYKRGITPKBSLLKALILAGADVGLG 300
 QY 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYTFITAGKPLKISLWSPAPASTTA 360
 DB 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYTFITAGKPLKISLWSPAPASTTA 360
 QY 361 SVTLVNDLIDVITAPNCTRVGNDFSAPFNNMDGRNNVENVTINSPOSGTYITIEVQAN 420
 DB 361 SVTLVNDLIDVITAPNCTRVGNDFSAPFNNMDGRNNVENVTINSPOSGTYITIEVQAN 420
 QY 421 VPVGPONFSLAIYN 434
 DB 421 VPVGPONFSLAIYN 434
 RESULT 12
 AAM50084
 ID AAM50084 standard; protein; 433 AA.
 XX AC AAM50084;
 XX 12-AUG-2002 (first entry)
 XX *Bacillus* sp SD-521 (FERM BP-11162) alkaline protease protein fragment.
 XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX *Bacillus* sp.
 XX EP1209233-A2.
 XX 29-MAY-2002.
 XX 22-NOV-2001; 2001EP-00127851.
 XX 22-NOV-2000; 2000JP-00355166.
 XX 12-APR-2001; 2001JP-00114048.
 XX (KAOS) KAO CORP.
 XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX WPI; 2002-437518/47.
 DR New modified alkaline proteases useful in detergent compositions.
 PT Claim 5; Page 16-18; 25pp; English.
 XX This invention describes novel *Bacillus* sp. alkaline proteases useful in

CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease SD-521 from the
 CC *Bacillus* sp strain SD-521 (FERM BP-11162) described in the method of the
 CC invention

XX Sequence 433 AA;

Query Match 89.6%; Score 2010.5; DB 5; Length 433;
 Best Local Similarity 88.5%; Pred. No. 14e-138;
 Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRIN 60
 DB 1 NDVARGIVKADVAQNNYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRIN 60
 QY 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLPFQAFS 120
 DB 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLPFQAFS 119
 QY 121 AGARHTNSWGAAVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARHTNSWGAAPVNGAYTANSRQVDEYRNNDMTILFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENTLRPSFGSYADNINHYAOTSSRGPTKDGRIKPDVMAFGTYTILSARSLAPDSSE 240
 DB 180 TVGATENTLRPSFGSLADNPNHIAQSSRGATRDGRIKPDVTAFGTYTILSARSLAPDSSE 239
 QY 241 MANHDSKYAVMGTSNATPIVAGNVAQLREHFVKRGIIPKPSILKALIJAGADVGLGY 300
 DB 240 MANHDSKYAVMGTSNATPIVAGNVAQLREHFVKRGIIPKPSILKALIJAGADVGLGY 299
 QY 301 PNGNQGWRVTLDKSLNVAIVNESSALSTSQKATYTFATAGKPKISLWSDAPASTTA 360
 DB 300 PNGDQGWGRVTLDKSLNVAIVNEATLTATGQKATYSFQAGKPKISLWTDAPASTTA 359
 QY 361 SYTLVNDLDTVITAPNGTRYGNDPFAFPDNNMDGNNVENYFINSPOGTYTIEVOAYN 420
 DB 360 SYTLVNDLDTVITAPNGQKTVGNDPFPYDNNMDGNNVENYFINAPOGTYTIEVOAYN 419
 QY 421 VPVGPQNFSLAIYN 434
 DB 420 VPSGPQRFSLAIYH 433

RESULT 13

AA050082 standard; protein; 433 AA.

XX AAM50082;

DT 12-AUG-2002 (first entry)

DE *Bacillus* sp D6-(FERM P1592) alkaline protease protein fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX *Bacillus* sp.

XX EP1209233-A2.

XX 29-MAY-2002.

XX 22-NOV-2001; 2001EP-00127851.

XX 22-NOV-2000; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.

XX (KAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N,
 PI Okuda M, Sasaki K;

XX WPI; 2002-437518/47.

PT New modified alkaline proteases useful in detergent compositions.

XX Claim 5; Page 13-15; 25pp; English.

XX This invention describes novel *Bacillus* sp. alkaline proteases useful in

CC detergent compositions, especially in laundry, bleaching or automatic

CC dishwasher detergents. The novel proteases have an increased detergency &

CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

CC sequence represents a fragment of the alkaline protease F-1 from *Bacillus*

CC sp strain D6-(FERM-P1592) described in the method of the invention

XX Sequence 433 AA;

Query Match 89.4%; Score 2006.5; DB 5; Length 433;
 Best Local Similarity 88.2%; Pred. No. 2.8e-138;
 Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRIN 60
 DB 1 NDVARGIVKADVAQNNYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRIN 60
 QY 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLPFQAFS 120
 DB 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLPFQAFS 119
 QY 121 AGARHTNSWGAAVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARHTNSWGAAPVNGAYTANSRQVDEYRNNDMTILFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENTLRPSFGSYADNINHYAOTSSRGPTKDGRIKPDVMAFGTYTILSARSLAPDSSE 240
 DB 180 TVGATENTLRPSFGSLADNPNHIAQSSRGATRDGRIKPDVTAFGTYTILSARSLAPDSSE 239
 QY 241 MANHDSKYAVMGTSNATPIVAGNVAQLREHFVKRGIIPKPSILKALIJAGADVGLGY 300
 DB 240 MANHDSKYAVMGTSNATPIVAGNVAQLREHFVKRGIIPKPSILKALIJAGADVGLGY 299
 QY 301 PNGNQGWRVTLDKSLNVAIVNESSALSTSQKATYTFATAGKPKISLWSDAPASTTA 360
 DB 300 PNGDQGWGRVTLDKSLNVAIVNEATLTATGQKATYSFQAGKPKISLWTDAPASTTA 359
 QY 361 SYTLVNDLDTVITAPNGTRYGNDPFAFPDNNMDGNNVENYFINSPOGTYTIEVOAYN 420
 DB 360 SYTLVNDLDTVITAPNGQKTVGNDPFPYDNNMDGNNVENYFINAPOGTYTIEVOAYN 419
 QY 421 VPVGPQNFSLAIYN 434
 DB 420 VPSGPQRFSLAIYH 433

RESULT 14

AA089548 standard; protein; 636 AA.

XX AAM89548;

DT 12-APR-1999 (first entry)

DE *Bacillus* sp. alkaline protease Y.

XX Alkaline protease Y; detergent; surfactant; leather processing;

XX debittering; flavour.

XX *Bacillus* sp.

XX WO9856927-A2.

XX 17-DEC-1998.

XX 09-JUN-1998; 98WO-US012005.

XX 12-JUN-1997; 9705-00873479.
 PR (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Sloma A, Christianson L;
 XX
 DR WPI, 1999-080908/07.
 XX
 PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing.
 XX
 PS Claim 3; Page 55-56; 77pp; English.
 XX
 CC This is the amino acid sequence of a *Bacillus* sp. alkaline protease Y
 CC that is said to have good alkali and surfactant resistance and improved
 CC detergency. It shows 77% identity to a newly isolated protease (see
 CC AAM89547) of *Bacillus* sp. JF170 (NCIB 12513). The invention provides
 CC vectors, recombinant host cells and methods for the recombinant
 CC production of such proteases. The protease are used in laundry and
 CC dishwashing detergents, for institutional and industrial cleaning, and
 CC for leather processing, as well as for debittering and enhancing the
 CC degree of hydrolysis of protein hydrolysates, for flavour development
 CC through hydrolysis of proteins, degradation of undesired peptides and in
 CC enzymatic synthesis of peptides. They have enhanced stability towards
 CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
 CC peroxy type. The invention also provides mutant cells in which the
 CC protease activity is diminished. Such cells can be used for the
 CC production of heterologous recombinant proteins
 XX
 SQ Sequence 636 AA;
 Query Match 89.3%; Score 2004.5; DB 2; Length 636;
 Best Local Similarity 88.2%; Pred. No. 6.5e-138;
 Matches 383; Conservative 27; Mismatches 23; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQSSGLYGGQGIYAVADTGLDTGRNDSMHEAFRGKITALYALGRIN 60
 DB 204 NDVARGIVKADVAQNNNGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRIN 263
 QY 61 NANDNGHGTIVAGSVLGNATNKMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAFS 120
 DB 264 NASDPNGHGTIVAGSVLGN-ALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAMN 322
 QY 121 AGARIHNSWGAAYNGAYTTDSRNVDYRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 180
 DB 323 AGARIHNSWGAAYNGAYTTDSRNVDYRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 382
 QY 181 TVGATENTLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
 DB 383 TVGATENTLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 442
 QY 241 WANHDSKYAVMGTSMTPTIVAGNVAQLRHFVYKRGITPKPSILKAALJAGADVGLGY 300
 DB 443 WANHDSKYAVMGTSMTPTIVAGNVAQLRHFVYKRGITPKPSILKAALJAGADVGLGY 502
 QY 301 PNGOGGKRVTLDKSLNVAAYVNESSALSTOKATYTFETAAGKELISLWSPAPASTA 360
 DB 503 PNGOGGKRVTLDKSLNVAAYVNESSALSTOKATYTFETAAGKELISLWSPAPASTA 562
 QY 361 SVTLVNDLVLITAPNGTRVYGNDFSAFPDNNMDGRNNVENVINSPOSGTYTIEVOAYN 420
 DB 563 SVTLVNDLVLITAPNGTRVYGNDFSAFPDNNMDGRNNVENVINSPOSGTYTIEVOAYN 622
 QY 421 VPVGPONFSLAIYN 434
 DB 623 VPVGPONFSLAIYN 636

AC AAM50083;
 XX 12-AUG-2002 (first entry)
 DT
 XX
 DE *Bacillus* sp. Y-(FERM BP-1029) alkaline protease protein fragment.
 XX
 KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS *Bacillus* sp.
 XX
 PN EP1209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-00127851.
 XX
 PR 22-NOV-2000; 2000JP-00355166.
 PR 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAO) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Arai H, Sumitomo N;
 PI Okuda M, Saeki K;
 DR WPI, 2002-437518/47.
 PT New modified alkaline proteases useful in detergent compositions.
 PS Claim 5; Page 15-16; 25pp; English.
 XX
 CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease Ya from *Bacillus*
 CC sp. strain Y-(FERM BP-1029) described in the method of the invention
 XX
 SQ Sequence 433 AA;
 Query Match 89.1%; Score 1999.5; DB 5; Length 433;
 Best Local Similarity 88.0%; Pred. No. 9e-138;
 Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQSSGLYGGQGIYAVADTGLDTGRNDSMHEAFRGKITALYALGRIN 60
 DB 1 NDVARGIVKADVAQNNNGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRIN 60
 QY 61 NANDNGHGTIVAGSVLGNATNKMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAFS 120
 DB 61 NASDPNGHGTIVAGSVLGN-ALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAMN 119
 QY 121 AGARIHNSWGAAYNGAYTTDSRNVDYRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 180
 DB 120 AGARIHNSWGAAYNGAYTTDSRNVDYRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 179
 QY 181 TVGATENTLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
 DB 180 TVGATENTLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 229
 QY 241 WANHDSKYAVMGTSMTPTIVAGNVAQLRHFVYKRGITPKPSILKAALJAGADVGLGY 300
 DB 240 WANHDSKYAVMGTSMTPTIVAGNVAQLRHFVYKRGITPKPSILKAALJAGADVGLGY 299
 QY 301 PNGOGGKRVTLDKSLNVAAYVNESSALSTOKATYTFETAAGKELISLWSPAPASTA 360
 DB 300 PNGOGGKRVTLDKSLNVAAYVNESSALSTOKATYTFETAAGKELISLWSPAPASTA 359
 QY 361 SVTLVNDLVLITAPNGTRVYGNDFSAFPDNNMDGRNNVENVINSPOSGTYTIEVOAYN 420
 DB 360 SVTLVNDLVLITAPNGTRVYGNDFSAFPDNNMDGRNNVENVINSPOSGTYTIEVOAYN 419
 QY 421 VPVGPONFSLAIYN 434

Sat Apr 3 17:19:58 2004

us-09-985-689a-2.rag

Page 11

DB 420 VPSGPRFSLAIVH 433

Search completed: March 31, 2004, 16:04:29
Job time : 49.2392 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 31, 2004, 15:59:39 ; Search time 14.0215 Seconds
(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689A-2

Perfect score: 2245

Sequence: 1 NDVARGIVKADVAQSSYGLY.....EVQAINVPFGPQNSLAIVN 434

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database: Issued Patents AA:*

1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/CTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2237	99.6	639	4	US-09-509-814A-4
2	2197	97.9	640	4	US-09-509-814A-8
3	2181	97.6	640	4	US-09-509-814A-6
4	2155	96.0	639	4	US-09-509-814A-1
5	2155	96.0	640	4	US-09-509-814A-2
6	2118.5	94.4	431	2	US-08-873-479-42
7	1998.5	89.0	433	4	US-09-104-623A-4
8	1998.5	89.0	433	4	US-09-019-533-4
9	1998.5	89.0	433	4	US-09-338-746-4
10	1998.5	89.0	635	2	US-08-873-479-43
11	1579.5	70.4	345	4	US-09-512-251A-10
12	1579.5	70.4	345	4	US-09-515-150A-10
13	1579.5	70.4	345	4	US-09-196-281-13
14	1579.5	70.4	345	4	US-08-894-818B-1
15	1579.5	70.4	345	4	US-08-894-818B-1
16	1579.5	70.4	345	4	US-09-445-472-12
17	1579.5	70.4	345	4	US-09-445-472-1
18	1579.5	70.4	345	4	US-08-894-818B-3
19	1579.5	70.4	345	4	US-08-894-818B-3
20	1579.5	70.4	345	4	US-08-894-818B-3
21	1579.5	70.4	345	4	US-09-445-472-16
22	1579.5	70.4	345	4	US-08-894-818B-5
23	1579.5	70.4	345	4	US-09-000-016-7
24	1579.5	70.4	345	4	US-09-514-340-7
25	1579.5	70.4	345	4	US-09-000-016-4
26	1579.5	70.4	345	4	US-09-514-340-4
27	1579.5	70.4	345	4	US-09-000-016-2
28	1579.5	70.4	345	4	US-09-514-340-2

28	304.5	13.6	903	1	US-08-750-532-1	Sequence 1, Appli
29	304.5	13.6	1398	1	US-08-750-532-9	Sequence 9, Appli
30	304.5	13.6	1398	3	US-08-894-818B-8	Sequence 6, Appli
31	304.5	13.6	1398	4	US-09-445-472-6	Sequence 18, Appli
32	296.5	13.2	237	1	US-08-750-532-18	Sequence 44, Appli
33	285.5	12.7	418	2	US-08-873-479-44	Sequence 2, Appli
34	273	12.2	418	4	US-09-966-921A-2	Sequence 8, Appli
35	260.5	11.6	397	4	US-08-328-352-7533	Sequence 1, Appli
36	259.5	11.6	275	1	US-08-431-387-1	Sequence 8, Appli
37	256	11.4	280	1	US-08-434-255-8	Sequence 8, Appli
38	256	11.4	280	1	US-08-459-967-8	Sequence 8, Appli
39	256	11.4	280	1	US-08-460-327-8	Sequence 8, Appli
40	256	11.4	280	1	US-08-459-871-8	Sequence 8, Appli
41	256	11.4	280	3	US-09-024-532-2	Sequence 2, Appli
42	256	11.4	280	4	US-09-104-623A-2	Sequence 2, Appli
43	256	11.4	280	4	US-09-019-532-2	Sequence 2, Appli
44	256	11.4	280	4	US-09-417-359A-2	Sequence 2, Appli
45	256	11.4	280	4	US-09-705-185-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1	US-09-509-814A-4	
Sequence 4, Application US/09509814A		
Patent No. 6376227		
GENERAL INFORMATION:		
APPLICANT: TAKAWA, MIKIO		
APPLICANT: OKUDA, MITSUYOSHI		
APPLICANT: SAKI, KATSUHIKA		
APPLICANT: KUBOTA, HIROMI		
APPLICANT: HITOMI, JUN		
APPLICANT: KAGEYAMA, YASUSHI		
APPLICANT: SHIKATA, SHITSUMI		
APPLICANT: NOMURA, MASAFUMI		
TITLE OF INVENTION: ALKALINE PROTEASE		
FILE REFERENCE: 0327-0832-OPCT		
CURRENT APPLICATION NUMBER: US/09/509,814A		
CURRENT FILING DATE: 2000-04-06		
PRIOR APPLICATION NUMBER: PCT/JP98/04528		
PRIOR FILING DATE: 1998-10-07		
PRIOR APPLICATION NUMBER: JP 9-274570		
PRIOR FILING DATE: 1997-06-08		
NUMBER OF SEQ ID NOS: 24		
SOFTWARE: Patent version 3.0		
SEQ ID NO 4		
LENGTH: 639		
TYPE: PRT		
ORGANISM: Bacillus sp.		
US-09-509-814A-4		
Query Match	99.6%	Score 2237; DB 4; Length 639;
Best Local Similarity	99.8%	Pred. No. 7.4e-175;
Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	NDVARGIVKADVAQSSYGLYGGQIIVAVADTGIDTGRNDSNHEAFRGKITALYALGRIN 60
DB	206	NDVARGIVKADVAQSSYGLYGGQIIVAVADTGIDTGRNDSNHEAFRGKITALYALGRIN 265
QY	61	NANDNGHGHVAVGSLGNGATKMGAPANLVFQSIMSSGGLGGLPENTLQTLFSGARS 120
DB	266	NANDNGHGHVAVGSLGNGATKMGAPANLVFQSIMSSGGLGGLPENTLQTLFSGARS 325
QY	121	AGARIHNSGAAVNGAYTTDSRNVDDYVRKDMTILFAAGNEGPGGTISAPGTAKNAI 180
DB	326	AGARIHNSGAAVNGAYTTDSRNVDDYVRKDMTILFAAGNERPFGGTISAPGTAKNAI 385
QY	181	TGATNTLRSPSSYADNININHAQSSRGPTDGRKIPVMAPGYTLARSISLADSSSF 240
DB	386	TGATNTLRSPSSYADNININHAQSSRGPTDGRKIPVMAPGYTLARSISLADSSSF 445
QY	241	WANHSKRAYAGGTSNATPIVAGNVAAQLREHFVKRNGITPKPSLLKALILAGAADVGLGY 300

Sat Apr 3 17:19:59 2004

us-09-985-689a-2.rai

Page 2

Db 446 MAHNDKYAAWGSTSWATPIVAGNVAQLREHFVYKRGITPKRSLLKALLINGADVGLG 505
 Qy 301 PNNGMGKGRITLIDKSLINVAIVNESSALSTSCKAATTFATAGKPKIKSLVMSDAPASTTA 360
 Db 506 PNNGMGKGRITLIDKSLINVAIVNESSALSTSCKAATTFATAGKPKIKSLVMSDAPASTTA 565
 Qy 361 SVTLVNDLDELITAPNGIRYVGDFAFPDNNMDGRNNVENVFINSPOSGYTIIEVQAYN 420
 Db 566 SVTLVNDLDELITAPNGIRYVGDFAFPDNNMDGRNNVENVFINSPOSGYTIIEVQAYN 625
 Qy 421 VPPGQNFSLAIYN 434
 Db 626 VPPGQNFSLAIYN 639

```

RESULT 2
US-09-509-814A-8
Sequence 8, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAHUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0833-00CT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-509-814A-8

```

Query	Match	97.9%;	Score	2197;	DB	4	Length	640;	
Best Local Similarity		96.8%;	Pred. No.	1.4e-17;					
Matches	420;	Conservative	13;	Mismatches	1;	Indels	0;	Gaps	0
QY	1	NDVARGIVADAAQSSYGLVGGCGIVAAVDGLDGTGRDSSMHEAFRGKITALVALGRTN	60						
Db	207	NDVARGIVADAAQSSYGLVGGCGIVAAVDGLDGTGRDSSMHEAFRGKITALVALGRTN	266						
QY	61	NAVDTNGHGHVAGSYLNGATNKMAPOANLVQSIINDSGGLGGLPSNLQTLFEOQAS	120						
Db	267	NAVDTNGHGHVAGSYLNGATNKMAPOANLVQSIINDSGGLGGLPSNLQTLFEOQAS	326						
QY	121	AGARIHTNSGGAIVNVAATYTTDSRANDVDYRKNDMTLFLAANEGRNGTISAGTKAKNI	180						
Db	327	AGARIHTNSGGAIVNVAATYTTDSRANDVDYRKNDMTLFLAANEGRNGTISAGTKAKNI	386						
QY	181	TVGATENLPPSPFGSYADININHAQFSSSGPTKDGRIKPDVAPGTYTLLSARSSLADSSR	240						
Db	387	TVGATENLPPSPFGSYADININHAQFSSSGPTKDGRIKPDVAPGTYTLLSARSSLADSSR	446						
QY	241	WAMHDEKAYMGSTSNATPIVAGNVAQREHFVYKRGITTPPSILKALILAGADYGLG	300						
Db	447	WAMHDEKAYMGSTSNATPIVAGNVAQREHFVYKRGITTPPSILKALILAGADYGLG	506						
QY	301	PNGNQMGKATLDDKSLINVAIVNESSALSTQKATYFATGKGLKSLVWSAPASTA	360						
Db	507	PNGNQMGKATLDDKSLINVAIVNESSALSTQKATYFATGKGLKSLVWSAPASTA	566						

Qy	361	SVTLVNDLDELVTITAPKGTGYVGNDSAPFPDNNMVGANNVEVFINSPQSGYTTIEVQAYN	440
		: : :	
Db	567	SVTLVNDLDELVTITAPKGTGYVGNDETSPYNNDMGRNNVEVFINAPQSGYTTIEVQAYN	626
Qy	421	VPVGPQNFSLATYN	434
Db	627	VPVGPQNFSLATYN	640

```

RESULT 3
US-09-509-814A-6
Sequence 6, Application US/03509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKEMI, KATSUSHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASATSUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-05CT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: PCT/JP96/04528
PRIORITY FILING DATE: 1998-10-07
PRIORITY APPLICATION NUMBER: JP-9-274570
PRIORITY FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-509-814A-6

```

Query Matchc	97.6%	Score 2191	DB 4	Length 640
Best Local Similarity	96.5%	Pred. No. 4.3e-171		
Matches 419	Conservative 13	Mismatches 2	Indels 0	Gaps 0
QY	1	NDVARGIVKADVAAGSSYGLYGQGGIIVAVADTGLDTGRSDSMHEAFRGKITALVALGRTN	60	
Db	207	NDVARGIVADVAOOSYGLYGQGGIIVAVADTGLDTGRSDSMHEAFRGKITALVALGRTN	266	
QY	61	NANDTNGGHTHVAGSVLGNCA.TNKGMA.PQANLVFQSI.MOSGGGLGSLPENTLTLP.SQAFS	120	
Db	267	NANDTNGGHTHVAGSVLGNCS.TNKGMA.PQANLVFQSI.MOSGGGLGSLPENTLTLP.SQAFS	326	
QY	121	AGARIHTNSWGA.VAVGAYTTDS.RRVDDYVRKNDWTIL.PAAGNEGPGGCTISAPGTAKAI	180	
Db	327	AGARIHTNSWGA.VAVGAYTTDS.RRVDDYVRKNDWTIL.PAAGNEGPGGCTISAPGTAKAI	386	
QY	181	TYGATEHNLR.PSPGSADININVAOFFSRGRTQGR.IKPDVMAFGTYILSABSLAPDSSF	240	
Db	387	TYGATEHNLR.PSPGSADININVAOFFSRGRTQGR.IKPDVMAFGTYILSABSLAPDSSF	446	
QY	241	MANHDSKAYVMGGTSMATPIVAGVVAQLRHF.PYKNGITPKPSLTPAALIAAADVGGLGY	300	
Db	447	MANHDSKAYVMGGTSMATPIVAGVVAQLRHF.PYKNGITPKPSLTPAALIAAADIGLGY	506	
QY	301	PNGNQCKGRVTLDRSLNLAAYNESSALSTGQKATYTFPTATAGPPLKLSLWMDAPASTTA	360	
Db	507	PNGNQCKGRVTLDRSLNLAAYNESSALSTGQKATYTFPTATAGPPLKLSLWMDAPASTTA	566	
QY	361	SVTLVWDLDELVTAPNGRIRYAGNFSAP.FPMNMDGRNNEVYINS.PQSGTYTIEVOAYN	420	
Db	567	SVTLVWDLDELVTAPNGRQYVGNDFTEP.PYDNDMDGRNNEVYINAPQSGTYTIEVOAYN	626	
QY	421	VVPGPQNTSLAIYN 434		
Db	627	VVPGPQNTSLAIYN 640		

```
RESULT 4
US-09-509-814A-1
Sequence 1, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
```

```
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 96.0%; Score 2155; DB 4; Length 639;
Best Local Similarity 96.3%; Pred. No. 3.8e-168;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAGQSSYGLYGQGIYAVADTG:DTGRNDSMHEARFKITLALYALGRIN 60
DB 206 NDVARGIVKADVAGQSSYGLYGQGIYAVADTG:DTGRNDSMHEARFKITLALYALGRIN 265
QY 61 NANDTNGHGTAVAGSVLGNATNKMAPOANLVFQSIMSSGGLGGLPSNLQTLFQOAFS 120
DB 266 NANDTNGHGTAVAGSVLGNATNKMAPOANLVFQSIMSSGGLGGLPSNLQTLFQOAFS 325
QY 121 AGARHTNSMGAANVGYTDSRNVDYVRKNDMTLLFAAGNGEPNGGTTSAFGTAKNAI 180
DB 326 AGARHTNSMGAANVGYTDSRNVDYVRKNDMTLLFAAGNGEPNGGTTSAFGTAKNAI 385
QY 181 TVGATENLRPSFGSYADININHVAFSSRGPTKDKRIKPDVMAEGTYILSARSSLAPDSF 240
DB 386 TVGATENLRPSFGSYADININHVAFSSRGPTKDKRIKPDVMAEGTYILSARSSLAPDSF 445
QY 241 WANDSKYAYVGGTSMATPIYAGVADLRHFYVNGRITPKPBLAALAGADVGLGY 300
DB 446 WANDSKYAYVGGTSMATPIYAGVADLRHFYVNGRITPKPBLAALAGADVGLGY 505
QY 301 PNGNGGGRVTLDSLVAAVYNESALSTSQKATYFTATAGSPKLISLVMSAPASTA 360
DB 506 PNGNGGGRVTLDSLVAAVYNESALSTSQKATYFTATAGSPKLISLVMSAPASTA 565
QY 361 SVTLVNDLDVITAPNGTRYVGNDFSAFPNNNDGRRNVENVFINSPOSGTYTIEOVA 420
DB 566 SVTLVNDLDVITAPNGTRYVGNDFSAFPNNNDGRRNVENVFINSPOSGTYTIEOVA 625
QY 421 VPVGPOFSLAIYN 434
DB 626 VPVGPOFSLAIYN 639

RESULT 5
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2

LENGTH: 640
TYPE: prt
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (124)..(124)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc.feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2

Query Match 96.0%; Score 2155; DB 4; Length 640;
Best Local Similarity 96.3%; Pred. No. 3,8e-168;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 NDVARGIVADVAQSSYGLYGGQOIYAVADTGLDGRNNSWHEAFRGKITLTVLALGRN 60
DB 207 NDVARGIVADVAQSSYGLYGGQOIYAVADTGLDGRNNSWHEAFRGKITLTVLALGRN 266
QY 61 NANDINGHGHVAGSVLNGATNKGAPANLVFQSIIMDSGGGLPSNLQTLFSGAQS 120
DB 267 NANDINGHGHVAGSVLNGATNKGAPANLVFQSIIMDSGGGLPSNLQTLFSGAQS 326

QY 121 AGARITNSWGAANVGAATTTDSRVDDYVRKXDMTILFAAGNEGPNGTISAPGAKAI 180
DB 327 AGARITNSWGAANVGAATTTDSRVDDYVRKXDMTILFAAGNEGPNGTISAPGAKAI 386
QY 181 TVGATENLRPSFGSVADNINNVAAQFSSRPTDGRKIPDVAPGTIILSARSLAPDSSE 240
DB 387 TVGATENLRPSFGSVADNINNVAAQFSSRPTDGRKIPDVAPGTIILSARSLAPDSSE 446
QY 241 WANHDSKYVMGTSVATPIVAGNVAKLREHFVKNGITPKPSILKAALILAGAADVGLY 300
DB 447 WANHDSKYVMGTSVATPIVAGNVAKLREHFVKNGITPKPSILKAALILAGAADVGLY 506
QY 301 PNGNMGVATLTKSLNVAVYNESSALSTOKATYFTATACKPLKISLWSDAPASTTA 360
DB 507 PNGNMGVATLTKSLNVAVYNESSALSTOKATYFTATACKPLKISLWSDAPASTTA 566
QY 361 SVTLVNDLIVITAPRGTYGVNDFSPAPDNMDGNNTVENYFINSQSGTYTIEVOAYN 420
DB 567 SVTLVNDLIVITAPRGTYGVNDFSPAPDNMDGNNTVENYFINSQSGTYTIEVOAYN 626
QY 421 VPIVGPONFSLAIVN 434
DB 627 VPIVGPONFSLAIVN 640

RESULT 6
US-08-873-479-42
Sequence 42, Application US/08873479
Patent No. 5891701
GENERAL INFORMATION:
APPLICANT: Sloana, Alan
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
TITLE OF INVENTION: Having Protease Activity
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58917010 No. 5891701disk of No. 5891701th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,479
FILING DATE: 12-JUN-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Agitis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 641 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-873-479-42

Query Match 94.4%; Score 2118.5; DB 2; Length 641;
Best Local Similarity 93.3%; Pred. No. 3,7e-165;
Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;

```

QY      1 NDVARGIVKADVAQSSSYGLYGQQLVAVADTGLDPTGRNDSMHEAFRGKITLALGRIN 60
DB      209 NDVARGIVKADVAQNNPFGILGQQLVAVADTGLDPTGRNDSMHEAFRGKITLALGRIN 268
QY      61 NANDNGHGHVAGSVLGNATNKGNAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
DB      269 NANDPFGHGHVAGSVLGN-ATNKGNAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 327
QY      121 AGARHTNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNSGTISAPGIAKNAI 180
DB      328 AGARHTNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNSGTISAPGIAKNAI 387
QY      181 TVGATENLRPSFSGSYADNINHVAQFSRGPTRKGRIPKPDVMAPGTYILSARSSLAPDSSF 240
DB      388 TVGATENLRPSFSGSYADNINHVAQFSRGPTRKGRIPKPDVMAPGTYILSARSSLAPDSSF 447
QY      241 WANHDSKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIIAGADVGLGY 300
DB      448 WANHDSKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIIAGADVGLGY 507
QY      301 PNGNGGWRVTLDKSLNVAAYNNESSALSTOKATYFTATAGPKLISLWSDAPASTTA 360
DB      508 PNGNGGWRVTLDKSLNVAAYNNESSALSTOKATYFTATAGPKLISLWSDAPASTTA 567
QY      361 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB      568 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 627
QY      421 VPVGPQNFSLATVN 434
DB      628 VPVGPQNFSLATVN 641

RESULT 7
US-09-104-623a-4
; Sequence 4, Application US/09104623a
; Patent No. 6303752
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Fatum, Tine Muxoil
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Roggen, Erwin Judo
; TITLE OF INVENTION: A Modified Polypeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6303752a No. 6303752a of No. 6303752ch America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,623a
; FILING DATE: 25-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 5256,200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: Bacillus sp. X
US-09-104-623a-4
Query Match      89.0%; Score 1998.5; DB 4; Length 433;
Best Local Similarity 88.0%; Pred. No. 1,4e-155;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY      1 NDVARGIVKADVAQSSSYGLYGQQLVAVADTGLDPTGRNDSMHEAFRGKITLALGRIN 60
DB      1 NDVARGIVKADVAQNNPFGILGQQLVAVADTGLDPTGRNDSMHEAFRGKITLALGRIN 60
QY      61 NANDNGHGHVAGSVLGNATNKGNAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
DB      61 NANDPFGHGHVAGSVLGN-ATNKGNAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 119
QY      121 AGARHTNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNSGTISAPGIAKNAI 180
DB      120 AGARHTNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNSGTISAPGIAKNAI 179
QY      181 TVGATENLRPSFSGSYADNINHVAQFSRGPTRKGRIPKPDVMAPGTYILSARSSLAPDSSF 240
DB      180 TVGATENLRPSFSGSYADNINHVAQFSRGPTRKGRIPKPDVMAPGTYILSARSSLAPDSSF 239
QY      241 WANHDSKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIIAGADVGLGY 300
DB      240 WANHDSKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIIAGADVGLGY 299
QY      301 PNGNGGWRVTLDKSLNVAAYNNESSALSTOKATYFTATAGPKLISLWSDAPASTTA 360
DB      300 PNGNGGWRVTLDKSLNVAAYNNESSALSTOKATYFTATAGPKLISLWSDAPASTTA 359
QY      361 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB      360 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 419
QY      421 VPVGPQNFSLATVN 434
DB      420 VPVGPQNFSLATVN 433

RESULT 8
US-09-019-532-4
; Sequence 4, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Prent, Annette
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922,204-US
; CURRENT APPLICATION NUMBER: US/09/019,532B
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0038/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-019-532-4
Query Match      89.0%; Score 1998.5; DB 4; Length 433;
Best Local Similarity 88.0%; Pred. No. 1,4e-155;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
DB 1 NDVARGIVKADVAQNNYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
QY 61 NANDTGHGHTVAGSVLGNATKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSQAFS 120
DB 61 NADDPHGHGHTVAGSVLGN-ALNKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSQAWN 119
QY 121 AGARHITNSGAAVNGAYTTDSRNVDYVRKNDMTLFAAGNEGPNNGTISAPGTAKNAI 180
DB 120 AGARHITNSGAAVNGAYTTANSRQVDEYVANNMTVLFAAGNEGPNNGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSADNINNHVAQFSSRGPTKGRIRKPDVMAFGTYILSARSSILAPDSF 240
DB 180 TVGATENLRPSFGSADNPNHIAQFSSRGATRDGRIRKPDVTAEGFTILSARSSILAPDSF 239
QY 241 WANHSKRYAMGQTSMTAPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 300
DB 240 WANYSKYAMGQTSMTAPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 299
QY 301 PNGNQGGRVTLDKSLNVAAYVNESSALSTQKATYFTATAGKPKLISLVMSDAPASTTA 360
DB 300 PSQDQGWGRVTLDKSLNVAAYVNEATATATGQKATYFSFOAQGPKLISLVMTDAPASTTA 359
QY 361 SVTLVNDLDLVITAPNGTRYGNDFSAPEPNMNDGRNNVENVFINSQSGTYTIEVOAYN 420
DB 360 SVTLVNDLDLVITAPNGQKTVGNDFSYFDNMNDGRNNVENVFINAPQSGTYTIEVOAYN 419
QY 421 VPVGPNFSLAIYN 434
DB 420 VPSGPQRFSLAIYH 433

```

RESULT 9

```

US-09-338-746-4
/ Sequence 4, Application US/09338746
/ Patent No. 6635526
/ GENERAL INFORMATION:
/ APPLICANT: Deussen, Heinz-Cosel
/ APPLICANT: Olsen, Arne A.
/ APPLICANT: Fatum, Tine M.
/ APPLICANT: Roggen, Erwin L.
/ TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
/ FILE REFERENCE: 5619,200-US
/ CURRENT APPLICATION NUMBER: US/09/338,746
/ EARLIER FILING DATE: 1999-06-23
/ EARLIER APPLICATION NUMBER: PA 1998 00809
/ EARLIER FILING DATE: 1998-06-23
/ EARLIER APPLICATION NUMBER: 60/091,461
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 4
/ LENGTH: 433
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
US-09-338-746-4

```

```

Query Match 89.0%; Score 1998.5; DB 4; Length 433;
Best Local Similarity 88.0%; Pred. No. 1,4e-155;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
DB 1 NDVARGIVKADVAQNNYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
QY 61 NANDTGHGHTVAGSVLGNATKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSQAFS 120
DB 61 NADDPHGHGHTVAGSVLGN-ALNKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSQAWN 119
QY 121 AGARHITNSGAAVNGAYTTDSRNVDYVRKNDMTLFAAGNEGPNNGTISAPGTAKNAI 180

```

```

DB 120 AGARHITNSGAAVNGAYTTANSRQVDEYVANNMTVLFAAGNEGPNNGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSADNINNHVAQFSSRGPTKGRIRKPDVMAFGTYILSARSSILAPDSF 240
DB 180 TVGATENLRPSFGSADNPNHIAQFSSRGATRDGRIRKPDVTAEGFTILSARSSILAPDSF 239
QY 241 WANHSKRYAMGQTSMTAPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 300
DB 240 WANYSKYAMGQTSMTAPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 299
QY 301 PNGNQGGRVTLDKSLNVAAYVNESSALSTQKATYFTATAGKPKLISLVMSDAPASTTA 360
DB 300 PSQDQGWGRVTLDKSLNVAAYVNEATATATGQKATYFSFOAQGPKLISLVMTDAPASTTA 359
QY 361 SVTLVNDLDLVITAPNGTRYGNDFSAPEPNMNDGRNNVENVFINSQSGTYTIEVOAYN 420
DB 360 SVTLVNDLDLVITAPNGQKTVGNDFSYFDNMNDGRNNVENVFINAPQSGTYTIEVOAYN 419
QY 421 VPVGPNFSLAIYN 434
DB 420 VPSGPQRFSLAIYH 433

```

RESULT 10

```

US-08-873-479-43
/ Sequence 43, Application US/08873479
/ Patent No. 5891701
/ GENERAL INFORMATION:
/ APPLICANT: Sloma, Alan
/ APPLICANT: Lyne, Christianson
/ TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
/ TITLE OF INVENTION: Having Protease Activity
/ NUMBER OF SEQUENCES: 57
/ CORRESPONDENCE ADDRESSES:
/ ADDRESS: No. 58917010 No. 5891701disk of No. 5891701th America
/ STREET: 405 Lexington Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10174
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/873,479
/ FILING DATE: 12-JUN-1997
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Agtis, Cheryl H
/ REGISTRATION NUMBER: 34,086
/ REFERENCE/DOCKET NUMBER: 5251,000-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-867-0123
/ TELEFAX: 212-878-9655
/ TRIEX:
/ INFORMATION FOR SEQ ID NO: 43:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 635 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-873-479-43

```

```

Query Match 89.0%; Score 1998.5; DB 2; Length 635;
Best Local Similarity 88.0%; Pred. No. 2,4e-155;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
DB 203 NDVARGIVKADVAQNNYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 262

```

```

QY 61 NANTNGHGHVAVSGLNGATNKGAPOANLVFQSIMDSGGGLGGLPSNLQTLFSQAFS 120
Db 263 NADDPNGHGHVAVSGLNGATNKGAPOANLVFQSIMDSGGGLGGLPSNLQTLFSQAFS 321
QY 121 AGARIHTNSWGAANVNGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 180
Db 322 AGARIHTNSWGAAPNGAYTANSRQVDEYVANNMTILFAAGNPGNGGTTISAPGTAKNAI 381
QY 181 TVGATEHLRPSFGSYAANINHVAFSSRGPTKGRKPDVWAPGTITLSARSLAPDSGF 240
Db 382 TVGATEHLRPSFGSIADNPVHIAQFSSRGATRDGRKPDVWAPGTITLSARSLAPDSGF 441
QY 241 WANHSKXYAMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSILKALIAAGADVGLGY 300
Db 442 WANHSKXYAMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSILKALIAAGADVGLGY 501
QY 301 PNGNGGWRVTLDKSLNVAAYVNSSALSTQKATYFTATAGRPKILSLVMSAPASTA 360
Db 502 PNGNGGWRVTLDKSLNVAAYVNEATATATQKATYFTATAGRPKILSLVMSAPASTA 561
QY 361 SVTLVNDLDTITAPNGTRYVNDPSAPDNNDGNNVENVINSPOGTYTIEVOAYN 420
Db 562 SVTLVNDLDTITAPNGQXVNDPSYPDNMNDGNNVENVINSPOGTYTIEVOAYN 621
QY 421 VPVGPQNFSLATYN 434
Db 622 VPVGPQNFSLATYN 635

```

RESULT 11

US-09-512-251A-10

```

; Sequence 10, Application US/09512251A
; Patent No. 655355
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelson, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349,204-US
; CURRENT APPLICATION NUMBER: US/09/512,251A
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
; US-09-512-251A-10

```

Query Match 70.4%; Score 1579.5; DB 4; Length 345;

Best Local Similarity 95.6%; Pred. No. 2e-121; Indels 1; Gaps 1;

Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

```

QY 1 NDVARGIVKADVAAOSSGILYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 60
Db 29 NDVARGIVKADVAAQNNGFLYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 88
QY 61 NANTNGHGHVAVSGLNGATNKGAPOANLVFQSIMDSGGGLGGLPSNLQTLFSQAFS 120
Db 89 NADDPNGHGHVAVSGLNGATNKGAPOANLVFQSIMDSGGGLGGLPSNLQTLFSQAFS 147
QY 121 AGARIHTNSWGAANVNGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 180
Db 148 AGARIHTNSWGAAPNGAYTANSRQVDEYVANNMTILFAAGNPGNGGTTISAPGTAKNAI 207
QY 181 TVGATEHLRPSFGSYAANINHVAFSSRGPTKGRKPDVWAPGTITLSARSLAPDSGF 240
Db 208 TVGATEHLRPSFGSYADNPVHIAQFSSRGATRDGRKPDVWAPGTITLSARSLAPDSGF 267
QY 241 WANHSKXYAMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSILKALIAAGADVGLGY 300
Db 268 WANHSKXYAMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSILKALIAAGADVGLGY 327

```

```

QY 301 PNGNGGWRVTLDKSLNV 318
Db 328 PNGNGGWRVTLDKSLNV 345

```

RESULT 12

US-09-515-150A-10

```

; Sequence 10, Application US/09515150A
; Patent No. 6558938
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelson, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5348,204-US
; CURRENT APPLICATION NUMBER: US/09/515,150A
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
; US-09-515-150A-10

```

Query Match 70.4%; Score 1579.5; DB 4; Length 345;

Best Local Similarity 95.6%; Pred. No. 2e-121; Indels 1; Gaps 1;

Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

```

QY 1 NDVARGIVKADVAAOSSGILYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 60
Db 29 NDVARGIVKADVAAQNNGFLYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 88
QY 61 NANTNGHGHVAVSGLNGATNKGAPOANLVFQSIMDSGGGLGGLPSNLQTLFSQAFS 120
Db 89 NADDPNGHGHVAVSGLNGATNKGAPOANLVFQSIMDSGGGLGGLPSNLQTLFSQAFS 147
QY 121 AGARIHTNSWGAANVNGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 180
Db 148 AGARIHTNSWGAAPNGAYTANSRQVDEYVANNMTILFAAGNPGNGGTTISAPGTAKNAI 207
QY 181 TVGATEHLRPSFGSYAANINHVAFSSRGPTKGRKPDVWAPGTITLSARSLAPDSGF 240
Db 208 TVGATEHLRPSFGSYADNPVHIAQFSSRGATRDGRKPDVWAPGTITLSARSLAPDSGF 267
QY 241 WANHSKXYAMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSILKALIAAGADVGLGY 300
Db 268 WANHSKXYAMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSILKALIAAGADVGLGY 327
QY 301 PNGNGGWRVTLDKSLNV 318
Db 328 PNGNGGWRVTLDKSLNV 345

```

RESULT 13

US-09-196-281-13

```

; Sequence 13, Application US/09196281A
; Patent No. 6605458
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelson, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5435,200-US
; CURRENT APPLICATION NUMBER: US/09/196,281A
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 1332/97
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 13

```



```

/ LENGTH: 345
/ TYPE: PRT
/ ORGANISM: Bacillus
US-09-196-281-13

Query Match
Best Local Similarity 70.4%; Score 1579.5; DB 4; Length 345;
Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 NDVAGIVADYAGSSYGLYGQGIYAVADTGLDGRNDSNHEAFRGKITALLYALGRTN 60
DB 29 NDVAGIVADYAGSSYGLYGQGIYAVADTGLDGRNDSNHEAFRGKITALLYALGRTN 88
QY 61 NANDNGHGTHTVAGSVLGNATNKGAPOANLVFOSIMDSGGLGGLPSNLQTFPSQAFS 120
DB 89 NANDNGHGTHTVAGSVLGN-ATNKGAPOANLVFOSIMDSGGLGGLPSNLQTFPSQAFS 147
QY 121 AGARHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNGGTISAPGTAKXAI 180
DB 148 AGARHTNSWGAAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNGGTISAPGTAKXAI 207
QY 181 TTAGATENLRPSFGSYADNINNHVAFSSRGPTDGRIKPDVMAFGTYIIISASSSLAPDSF 240
DB 208 TTAGATENLRPSFGSYADNINNHVAFSSRGPTDGRIKPDVMAFGTYIIISASSSLAPDSF 267
QY 241 WANHDSKYAMGTSMATPIVAGNVAQLREHFVNKRGITPKPSLKAALLIAGADVGLGY 300
DB 268 WANHDSKYAMGTSMATPIVAGNVAQLREHFVNKRGITPKPSLKAALLIAGADVGLGY 327
QY 301 PNGNGMGRTLLDKSLNV 318
DB 328 PNGNGMGRTLLDKSLNV 345

RESULT 14
US-08-894-818B-1
/ Sequence 1, Application US/08894818B
/ Patent No. 6261822
/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: YAMAMOTO, Katsuhiko
/ APPLICANT: MITTA, Masamori
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: TSUNASAWA, Susumu
/ APPLICANT: KATO, Ikunoshin
/ TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Broadway and Neimark
/ STREET: 419 Seventh Street N.W., Ste. 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: United States of America
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/894,818B
/ FILING DATE: 20-MAY-1998
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP96/03253
/ FILING DATE: 07-NOV-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 333285/1995
/ FILING DATE: 12-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Browdy, Roger L.
/ REGISTRATION NUMBER: 25,618

```

```

/ REFERENCE/DOCKET NUMBER: TAKAKURA-1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-5197
/ TELEFAX: (202) 737-3528
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 659 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-894-818B-1

Query Match
Best Local Similarity 20.4%; Score 457.5; DB 3; Length 659;
Matches 136; Conservative 69; Mismatches 153; Indels 101; Gaps 17;

QY 8 VKADVAGSSYGLYGQGIYAVADTGLDGRNDSNHEAFRGKITALLY-ALGRTNANDTN 66
DB 145 IGADTVNLSIGYDGGVVAIVDTGIDAN-----HPDLKGVLTGWYDAVNGRSTPYDQ 198
QY 67 GGTHTVAGSVLGNATNK--GMAPOANLVFOSIM--DSSGGLGGLPSNLQTLFSGAFSA 121
DB 199 GGTHTVAGSVLGNATNK--GMAPOANLVFOSIM--DSSGGLGGLPSNLQTLFSGAFSA 258
QY 122 GARI-----HTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNGGT 169
DB 259 GIRVYNLSLGSQSSDGTBLSQAVNNAMDA-----GIIVCAAGNSGNTYT 306
QY 170 ISAPGTAKXAIITGATENLRPSFGSYADNINNHVAFSSRGPTDGRIKPDVMAFGTYIIS 229
DB 307 VGSPAAASKVITVGA-----VDSNDNIAFSSRGPTADGRLKEVVAAGVDIIA 355
QY 230 AASSIAPDSFFMANHDSKYAMGTSMATPIVAG--NVAQLREHFVNKRGITPK--PSLTK 286
DB 356 PRAS---GTSMGPINDIYRKAGTSMATPHVSGVGAALLIQA-----PSWTDXK 404
QY 287 ALIAGA-----ADVGLGYPNGMGRTLLDKSL--NVAYNSSALSTSKATY 335
DB 405 TALIETADIVAPKEIIDIAVGA-----GRVAVVAIKYDPAKLTFFGSVADKGSATH 457
QY 336 TPTATAGPKLKLIVMSDAPASTTASVTIVNDLVLTPNGTRYGNDFSAFDDNNMG 395
DB 458 TPDVSGATFVATILYND-----TGSDDLVLIVPENGME-----VDISYTA 498
QY 396 RNNVENYFINSFGSYTIEVQANVPVGPNSLAIVN 434
DB 499 YGFEKVGYYNPATGATVTVAVSYK--GAANYQVDVVS 534

RESULT 15
US-09-445-472-12
/ Sequence 12, Application US/09445472
/ Patent No. 6358726
/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: SHIMOTO, Tomoko
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: KATO, Ikunoshin
/ TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
/ FILE REFERENCE: TAKAKURA-6
/ CURRENT APPLICATION NUMBER: US/09/445,472
/ FILING DATE: 1999-12-06
/ PRIOR APPLICATION NUMBER: 151969/1997
/ PRIOR FILING DATE: 1997-06-10
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 12
/ LENGTH: 659
/ TYPE: PRT
/ ORGANISM: Thermococcus celer
US-09-445-472-12

```

Sat Apr 3 17:19:59 2004

us-09-985-689a-2.rai

Page 10

Query Match	20.4%;	Score 457.5;	DB 4;	Length 659;
Best Local Similarity	29.6%;	Pred. No. 3.4e-29;		
Matches 136;	Conservative 69;	Mismatches 153;	Indels 101;	Gaps 17;

```

QY      8 VKADVAOSSVGLVGQQOICIVADVDTGLDTRNDSSMHEAFPRGKITLYL-ALGRTNANDTN 66
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     145 IGAIDTYWNSLGDSGVSWVALVDITGDIA-----HPLDKRGVIGMDAVNGRSTPYDDQ 198
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      67 GHGHTVAGSVLGNCAATNK--GMAPCANLVFGSIN--DSSGGLGGLPBNLTQLFSQAFA 121
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     199 GHGHTVAGIACVAGTSVNSQYICVAPAKLVGVKVLGAAGSGSVFTIIAGVDMVVQNDRKY 258
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      122 GARI-----HTSMGAANVGAATTDSRVNVDYYARKNMOTILFPAANGPNPGGT 169
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     259 GIRVINLSLSSGSSSPGCTSLSCAANNADA-----GIWCVAAGNSGPMTYT 306
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      170 ISAPGTAKNAIYGATENLRPFSGSYADNIINHVAOFSSRGPTKGRIKEPDMAFGTILS 229
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     307 VGSPPAAASKVITYGA-----YBNDNLIASFFSRGPTAAGRKLPEVNAAGVDIILA 355
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      230 ARSSLAPDSFMANHSKAYMGCGTSMATPIYAG-NVAQLREHFYKONRGITPK--PSLLK 286
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     356 PRAS--GTMGTPINDYYTKASGTSMAFPHVSGGALLQAH-----PSWTPEPKVK 404
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      287 AALIAGA-----ADVGLGYPNGNQGWRTLLKSL--NVAVYNESSALISGCATY 335
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     405 TALIEETDIAPAKEIDIAYGA-----GRNNVKALKIKDYAKLTPTYSVADAKSGATH 457
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      336 TFTAATACKPLKISLVMSDAEPASTTASVTLVNDLDIVITAIPNGTRYVNGDFSAFPDNWDG 355
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     458 TFDSGATFTALTALYMD-----TGSSSIDLYLDYPNGNE-----VDYSYTA 498
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      396 RNNVENVFINSPOSGYTTIEVOAYNNVPVPOQNFSLAITN 434
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     499 YYPEEKYGYNPDPAGIMTVAKVSYK--GAANYQDVVAS 534

```

Search completed: March 31, 2004, 16:11:52
Job time : 15.0215 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comogen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:09:09 ; Search time 33.2177 Seconds

(without alignments)
3418.697 Million cell updates/sec

Title: US-09-985-689A-2

Perfect score: 2245
Sequence: 1 NDVARGIVKADVAQSSGYL.....EVQATNVGPFQNSLATVN 434Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2245	100.0	434	10	US-09-985-689A-2
2	2191	97.6	434	10	US-09-985-689A-1
3	2191	97.6	434	15	US-10-385-662-2
4	2130	94.9	434	10	US-09-985-689A-6
5	2118.5	94.4	433	10	US-09-985-689A-7
6	2010.5	89.6	433	10	US-09-985-689A-5
7	2006.5	89.4	433	10	US-09-985-689A-3
8	1999.5	89.1	433	10	US-09-985-689A-4
9	1579.5	70.4	345	14	US-10-336-324-10
10	1579.5	70.4	345	14	US-10-403-105-13
11	457.5	20.4	659	13	US-10-090-624-12
12	423	18.8	412	13	US-10-090-624-1
13	423	18.8	522	13	US-10-090-624-4
14	423	18.8	654	13	US-10-090-624-16
15	365	16.3	1079	14	US-10-112-488-39

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	360.5	345	343	308.5	308.5	304.5	281.5	277	273	270	266	256	256	254	253	252	251.5	250	250	247	247	247	247	246	246	246	246	246	246	246
	16.1	15.4	15.3	13.7	13.7	13.6	12.5	12.3	12.2	12.0	11.8	11.4	11.4	11.3	11.3	11.2	11.2	11.1	11.1	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
	1208	1139	1139	519	1975	1388	595	580	418	1101	663	280	397	368	271	271	379	271	271	271	271	271	271	271	269	269	269	269	269	269
	US-10-156-761-13251	US-10-156-761-10856	US-10-314-657-4	US-10-084-846A-114	US-10-084-846A-4	US-10-090-624-6	US-09-927-827-59	US-09-927-827-55	US-09-966-921A-2	US-10-156-761-13934	US-10-424-599-255271	US-10-209-812-2	US-09-779-334A-5	US-10-344-231-3	US-09-813-408-2	US-10-242-549-56	US-09-813-408-6	US-10-242-549-52	US-10-242-549-54	US-10-242-549-46	US-10-242-549-48	US-10-242-549-50	US-10-242-549-60	US-08-322-678-10	US-09-837-235-16	US-09-600-854B-6	US-09-575-139-1	US-09-576-414-8	US-09-736-116-49	US-10-075-907-1
	Sequence 13551, A	Sequence 10856, A	Sequence 4, App1	Sequence 114, App	Sequence 59, App1	Sequence 6, App1	Sequence 5, App1	Sequence 55, App1	Sequence 12934, A	Sequence 255271, A	Sequence 2, App1	Sequence 5, App1	Sequence 3, App1	Sequence 2, App1	Sequence 56, App1	Sequence 6, App1	Sequence 52, App1	Sequence 12934, A	Sequence 46, App1	Sequence 48, App1	Sequence 50, App1	Sequence 60, App1	Sequence 10, App1	Sequence 16, App1	Sequence 6, App1	Sequence 1, App1	Sequence 8, App1	Sequence 49, App1	Sequence 1, App1	

ALIGNMENTS

RESULT 1
US-09-985-689A-2
Sequence 2, Application US/09985689A
Publication No. US2003002235:AI
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SAITO, TSYOYUKI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKET, KATSUHIKA
TITLE OR INVENTION: Alkaline proteases
FILE REFERENCE: 215483USO
CURRENT APPLICATION NUMBER: US/09/985.689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-2
Query Match 100.0%; Score 2245; DB 10; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.4e-189;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSGYLGGQGIYAVADTGDDTGRNDSWHEAFRGKITLALAGRTN 60
DB 1 NDVARGIVKADVAQSSGYLGGQGIYAVADTGDDTGRNDSWHEAFRGKITLALAGRTN 60
QY 61 NANDTNGHGHVAGSVLGNAGATNKGMAPQANTVFOSSIMSSGGLGGLPENTLQTLFSGQAFS 120

```

Db      61 NANDTNGHGHVASSVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLTQTLFSGQAFS 120
Qy      121 AGARIHNTSMGAAVNGAYTTDSRVVDYVYAKNDMTILFAAGNEGPNCGTISAPGTAKNAI 180
Db      121 AGARIHNTSMGAAVNGAYTTDSRVVDYVYAKNDMTILFAAGNEGPNCGTISAPGTAKNAI 180
Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240
Qy      241 WANHSKYAVYGGGSMATPIVAGNVAQLREHFVNKRGITPKPSILKALLAGAADVGLGY 300
Db      241 WANHSKYAVYGGGSMATPIVAGNVAQLREHFVNKRGITPKPSILKALLAGAADVGLGY 300
Qy      301 PNGNQGWRVTLDKSLNVAAYNESSALSTQKATYFTATAGKPLKISLWSDAPASTTA 360
Db      301 PNGNQGWRVTLDKSLNVAAYNESSALSTQKATYFTATAGKPLKISLWSDAPASTTA 360
Qy      361 SVTLVNDLDVITAPNGRTRYVGNDFSAPEFNNMNGRNVENVFINSQSGTYITTEVOAYN 420
Db      361 SVTLVNDLDVITAPNGRTRYVGNDFSAPEFNNMNGRNVENVFINSQSGTYITTEVOAYN 420
Qy      421 VPVGPQNFSLAIYN 434
Db      421 VPVGPQNFSLAIYN 434

```

RESULT 2

```

US-09-985-689a-1
; Sequence 1, Application US/09985689a
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483U0
; CURRENT APPLICATION NUMBER: US/09/985,689a
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689a-1

```

```

Query Match      97.6%; Score 2191; DB 10; Length 434;
Best Local Similarity 96.5%; Pred. No. 4.7e-184;
Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Db      1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Qy      61 NANDTNGHGHVASSVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLTQTLFSGQAFS 120
Db      61 NANDTNGHGHVASSVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLTQTLFSGQAFS 120
Qy      121 AGARIHNTSMGAAVNGAYTTDSRVVDYVYAKNDMTILFAAGNEGPNCGTISAPGTAKNAI 180
Db      121 AGARIHNTSMGAAVNGAYTTDSRVVDYVYAKNDMTILFAAGNEGPNCGTISAPGTAKNAI 180
Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240

```

```

Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240
Qy      241 WANHSKYAVYGGGSMATPIVAGNVAQLREHFVNKRGITPKPSILKALLAGAADVGLGY 300
Db      241 WANHSKYAVYGGGSMATPIVAGNVAQLREHFVNKRGITPKPSILKALLAGAADVGLGY 300
Qy      301 PNGNQGWRVTLDKSLNVAAYNESSALSTQKATYFTATAGKPLKISLWSDAPASTTA 360
Db      301 PNGNQGWRVTLDKSLNVAAYNESSALSTQKATYFTATAGKPLKISLWSDAPASTTA 360
Qy      361 SVTLVNDLDVITAPNGRTRYVGNDFSAPEFNNMNGRNVENVFINSQSGTYITTEVOAYN 420
Db      361 SVTLVNDLDVITAPNGRTRYVGNDFSAPEFNNMNGRNVENVFINSQSGTYITTEVOAYN 420
Qy      421 VPVGPQNFSLAIYN 434
Db      421 VPVGPQNFSLAIYN 434

```

RESULT 3

```

US-10-385-662-2
; Sequence 2, Application US/10385662
; Publication No. US20040002432A1
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: SAITO, KAZUHIRO
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: IZAWA, YOSHIFUMI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KOBAYASHI, TOHRU
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: Alkaline protease
; FILE REFERENCE: 234936U0
; CURRENT APPLICATION NUMBER: US/10/385,662
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: JP 2002-081428
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: JP 2002-165987
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: JP 2002-304230
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: JP 2002-304231
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp. KSM-KP43
US-10-385-662-2

```

```

Query Match      97.6%; Score 2191; DB 15; Length 434;
Best Local Similarity 96.5%; Pred. No. 4.7e-184;
Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Db      1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Qy      61 NANDTNGHGHVASSVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLTQTLFSGQAFS 120
Db      61 NANDTNGHGHVASSVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLTQTLFSGQAFS 120
Qy      121 AGARIHNTSMGAAVNGAYTTDSRVVDYVYAKNDMTILFAAGNEGPNCGTISAPGTAKNAI 180
Db      121 AGARIHNTSMGAAVNGAYTTDSRVVDYVYAKNDMTILFAAGNEGPNCGTISAPGTAKNAI 180
Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIPDVMAFGTYILSARSSLAPDSSF 240

```

```

QY 241 WANDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVLGY 300
DB 241 WANDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVLGY 300
QY 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYFTTAGKPLKLSLWSDAPASTTA 360
DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYFTTAGKPLKLSLWSDAPASTTA 360
QY 361 SVTLVNDLDIVITAPNGTRIVGNDPFSAPFDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB 361 SVTLVNDLDIVITAPNGTRIVGNDPFSAPFDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
QY 421 VPVGPQNFSLAIYN 434
DB 421 VPVGPQNFSLAIYN 434

```

RESULT 4

```

US-09-985-689a-6
; Sequence 6, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 434
; TYPE: PRP
; ORGANISM: Bacillus sp.
US-09-985-689a-6

```

Query Match 94.8%; Score 2130; DB 10; Length 434;

Best Local Similarity 92.4%; Pred. No. 1, 1e-178; Indels 0; Gaps 0;

```

Matches 401; Conservative 24; Mismatches 9; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
QY 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIMDSGGIGLPSNLQTLFSGAYS 120
DB 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIMDSGGIGLPSNLQTLFSGAYS 120
QY 121 AGARIHNTSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNGGTISAPGTAKNAI 180
DB 121 AGARIHNTSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNGGTISAPGTAKNAI 180
QY 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTKGRIPDMAPGTITLSARSLAPDSGF 240
DB 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTKGRIPDMAPGTITLSARSLAPDSGF 240
QY 241 WANDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVLGY 300
DB 241 WANDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVLGY 300
QY 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYFTTAGKPLKLSLWSDAPASTTA 360
DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYFTTAGKPLKLSLWSDAPASTTA 360

```

```

DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYFTTAGKPLKLSLWSDAPASTTA 360
QY 361 SVTLVNDLDIVITAPNGTRIVGNDPFSAPFDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB 361 SVTLVNDLDIVITAPNGTRIVGNDPFSAPFDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
QY 421 VPVGPQNFSLAIYN 434
DB 421 VPVGPQNFSLAIYN 434

```

RESULT 5

```

US-09-985-689a-7
; Sequence 7, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 433
; TYPE: PRP
; ORGANISM: Bacillus sp.
US-09-985-689a-7

```

Query Match 94.4%; Score 2118.5; DB 10; Length 433;

Best Local Similarity 93.3%; Pred. No. 1, 1e-177; Indels 1; Gaps 1;

Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
QY 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIMDSGGIGLPSNLQTLFSGAYS 120
DB 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIMDSGGIGLPSNLQTLFSGAYS 119
QY 121 AGARIHNTSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNGGTISAPGTAKNAI 180
DB 121 AGARIHNTSWGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPNGGTISAPGTAKNAI 179
QY 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTKGRIPDMAPGTITLSARSLAPDSGF 240
DB 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTKGRIPDMAPGTITLSARSLAPDSGF 239
QY 241 WANDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVLGY 300
DB 241 WANDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVLGY 299
QY 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYFTTAGKPLKLSLWSDAPASTTA 360
DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYFTTAGKPLKLSLWSDAPASTTA 359
QY 361 SVTLVNDLDIVITAPNGTRIVGNDPFSAPFDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB 361 SVTLVNDLDIVITAPNGTRIVGNDPFSAPFDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 419
QY 421 VPVGPQNFSLAIYN 434
DB 421 VPVGPQNFSLAIYN 434

```

Db 420 VPSGPFQFSLAIHV 433

RESULT 6
US-09-985-689A-5
Sequence 5, Application US/09985689A
Publication No. US2003022351A1
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SAITO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985, 689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 433
TYPE: PR1
ORGANISM: Bacillus sp.
US-09-985-689A-5

Query Match 89.6%; Score 2010.5; DB 10; Length 433;
Best Local Similarity 88.5%; Pred. No. 3.6e-168;
Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;
Db 1 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
1 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
Qy 61 NANDNGHGHVAGSVLGNATNGKVAPOANLVFQSIIMDSGGIGLPSNLQTLFSGAQS 120
61 NANDPNHGHTHVAGSVLGN-ALNKGVAPOANLVFQSIIMDSGGIGLPSNLQTLFSGAQN 119
Db 121 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 180
121 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 180
Db 120 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 179
120 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 179
Qy 181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQGYTILSARSSLADDSF 240
181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQGYTILSARSSLADDSF 240
Db 180 TVGATEENLRPSFGSLADNPNHIAQFSSRGATRDGRIRKPDVTAQGTILSARSSLADDSF 239
180 TVGATEENLRPSFGSLADNPNHIAQFSSRGATRDGRIRKPDVTAQGTILSARSSLADDSF 239
Qy 241 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 300
241 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 300
Db 240 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 299
240 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 299
Qy 301 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 360
301 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 360
Db 300 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 359
300 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 359
Qy 361 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 420
361 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 420
Db 360 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 419
360 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 419
Qy 421 VPSGPFQFSLAIHV 434
421 VPSGPFQFSLAIHV 434
Db 420 VPSGPFQFSLAIHV 433
420 VPSGPFQFSLAIHV 433

RESULT 7
US-09-985-689A-3
Sequence 3, Application US/09985689A

Publication No. US2003022351A1
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SAITO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985, 689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 433
TYPE: PR1
ORGANISM: Bacillus sp.
US-09-985-689A-3

Query Match 89.4%; Score 2006.5; DB 10; Length 433;
Best Local Similarity 88.2%; Pred. No. 8e-168;
Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;
Db 1 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
1 NDVARGIVKADVAQSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
Qy 61 NANDNGHGHVAGSVLGNATNGKVAPOANLVFQSIIMDSGGIGLPSNLQTLFSGAQS 120
61 NANDPNHGHTHVAGSVLGN-ALNKGVAPOANLVFQSIIMDSGGIGLPSNLQTLFSGAQN 119
Db 61 NANDPNHGHTHVAGSVLGN-ALNKGVAPOANLVFQSIIMDSGGIGLPSNLQTLFSGAQN 119
121 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 180
121 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 180
Db 120 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 179
120 AGARIHTNSWGAANGAYTTDSRVNDYVRKNDMTILFAAGNPGNGGTISAPGTAKNAI 179
Qy 181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQGYTILSARSSLADDSF 240
181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQGYTILSARSSLADDSF 240
Db 180 TVGATEENLRPSFGSLADNPNHIAQFSSRGATRDGRIRKPDVTAQGTILSARSSLADDSF 239
180 TVGATEENLRPSFGSLADNPNHIAQFSSRGATRDGRIRKPDVTAQGTILSARSSLADDSF 239
Qy 241 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 300
241 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 300
Db 240 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 299
240 WANHSKXAYMGTSMTPIVAGNVAQLREHFVKNRGITPKPSILKALILAGADVGLGY 299
Qy 301 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 360
301 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 360
Db 300 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 359
300 PNGQGWGRVTLDKSLNVAAYNESSALSTSOKATYFTATAGAPLKLISLWSDAPASTTA 359
Qy 361 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 420
361 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 420
Db 360 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 419
360 SVTLVNDLIDLVTAPNGRIVGNDFSAPDNNDGRNNVENVFINSQSGTYTIEVOAYN 419
Qy 421 VPSGPFQFSLAIHV 434
421 VPSGPFQFSLAIHV 434
Db 420 VPSGPFQFSLAIHV 433
420 VPSGPFQFSLAIHV 433

RESULT 8
US-09-985-689A-4
Sequence 4, Application US/09985689A
Publication No. US2003022351A1
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SAITO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI

```

; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
; US-09-985-689A-4

```

```

Query Match      89.1%; Score 1999.5; DB 10; Length 433;
Best Local Similarity 86.0%; Pred. No. 3.3e-167;
Matches 362; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNYGLYGQGVYAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDTNGHGHVAGSVLGNATNKGMAPQANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 120
DB 61 NANDTNGHGHVAGSVLGN-ALNKGMAPQANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 119
QY 121 AGARHTNSWGAANGAYTTDSRNVDDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 180
DB 120 AGARHTNSWGAANGAYTTDSRNVDDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 179
QY 181 TVGATENTLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAFGTYILSARSSLADPDSF 240
DB 180 TVGATENTLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAFGTYILSARSSLADPDSF 239
QY 241 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSILKAALIIAGADVGLGF 300
DB 240 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSILKAALIIAGADVGLGF 299
QY 301 PNGNQGWGRTYLDKSLNVAIVNVESSALSTSQATYITFTAGKPKLISLWSDAPASTTA 360
DB 300 PNGNQGWGRTYLDKSLNVAIVNVESSALSTSQATYITFTAGKPKLISLWSDAPASTTA 359
QY 361 SVTLVVDLVLVITAPNGTRYVGNDFSAFPDNNMDDGNVNEVFINSPOSGTYTIEVOAYN 420
DB 360 SVTLVVDLVLVITAPNGTRYVGNDFSAFPDNNMDDGNVNEVFINSPOSGTYTIEVOAYN 419
QY 421 VPVGPQNFSLAIYN 434
DB 420 VPEGPQRFSLAIYH 433

```

```

RESULT 9
US-10-336-324-10
; Sequence 10, Application US/10336324
; Publication No. US20030176304A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349,204-US
; CURRENT APPLICATION NUMBER: US/10/336,324
; PRIOR FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: US/09/5,2,251A
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
; US-10-336-324-10

```

```

Query Match      70.4%; Score 1579.5; DB 14; Length 345;
Best Local Similarity 95.6%; Pred. No. 2.1e-130;
Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 29 NDVARGIVKADVAQNNYGLYGQGVYAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 88
QY 61 NANDTNGHGHVAGSVLGNATNKGMAPQANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 120
DB 89 NANDTNGHGHVAGSVLGN-ATNKGMAPQANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 147
QY 121 AGARHTNSWGAANGAYTTDSRNVDDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 180
DB 148 AGARHTNSWGAANGAYTTDSRNVDDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 207
QY 208 TVGATENTLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAFGTYILSARSSLADPDSF 267
DB 241 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSILKAALIIAGADVGLGF 300
QY 268 WANHDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSILKAALIIAGADVGLGF 327
QY 301 PNGNQGWGRTYLDKSLNV 318
DB 328 PNGNQGWGRTYLDKSLNV 345

```

```

RESULT 10
US-10-403-105-13
; Sequence 13, Application US/10403105
; Publication No. US20030180933A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants And Compositions
; FILE REFERENCE: 5435,200-US
; CURRENT APPLICATION NUMBER: US/10/403,105
; CURRENT FILING DATE: 2003-03-31
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: US/09/196,281A
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1352/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
; US-10-403-105-13

```

```

Query Match      70.4%; Score 1579.5; DB 14; Length 345;
Best Local Similarity 95.6%; Pred. No. 2.1e-130;
Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 29 NDVARGIVKADVAQNNYGLYGQGVYAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 88
QY 61 NANDTNGHGHVAGSVLGNATNKGMAPQANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 120
DB 89 NANDTNGHGHVAGSVLGN-ATNKGMAPQANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 147
QY 121 AGARHTNSWGAANGAYTTDSRNVDDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 180

```

Sat Apr 3 17:19:59 2004

us-09-985-689a-2.rabp

Page 6

Db 148 AGARHITNSGAFVNGAYTTDSRNVDDYRKNDMTLLFAAGNEGSGGTISAPGTAKNAI 207
Qy 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVMAAGTYILSARSLAPDSF 240
Db 208 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVMAAGTYILSARSLAPDSF 267
Qy 241 WANHSKYAYMGSTVATPIVAGNVAQLREHFVNKGITPKPSLLKALIAAGADVIGY 300
Db 268 WANHSKYAYMGSTVATPIVAGNVAQLREHFVNKGITPKPSLLKALIAAGADVIGY 327
Qy 301 PNGNGGRTYLLDKSLNV 318
Db 328 PNGNGGRTYLLDKSLNV 345

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 20.4%; Score 457.5; DB 13; Length 659;
Best Local Similarity 29.6%; Pred. No. 2,7e-31;
Matches 136; Conservative 69; Mismatches 153; Indels 101; Gaps 17;

Qy 8 VKADVAQSYGLYGGQIVAVADTGLDTRNDSMEAFRGKITALY-ALGRTNANDN 66
Db 145 IGAULTVANSLSGYDSSGVVAIVDTGITAN-----HEDLKGRKTYGMTDAVNGSTYDDQ 198
Qy 67 GHGTHVAGSVLGNATNK--GMAPOANLVFQSIM--DSSGGLGGLPSNLQTLFSAQFA 121
Db 199 GHGTHVAGSVLGNATNK--GMAPOANLVFQSIM--DSSGGLGGLPSNLQTLFSAQFA 258
Qy 122 GARI-----HTNSGGAAYNGAYTTDSRNVDDYRKNDMTLLFAAGNEGNGGT 169
Db 259 GIRIYNISLSSGSSDGTDSLQAVNNAMDA-----GIYCVAAAGSGERTYI 306
Qy 170 ISAGTKAKNAITVATENLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVMAAGTYIL 229
Db 307 VGSFPAASKVITVGA-----VDSNDNLSFSSRGPTADGRIRKPEVVAPEGVDIA 355
Qy 230 ARSLADSSFWANHSKYAYMGSTVATPIVAG--NVAQLREHFVNKGITPK--PSLLK 286
Db 356 PRAS--GTSWGTINDYTYKASGTSMATPHVSGVGLIQAQ-----PSWTPDKYK 404
Qy 287 AALIGA-----ADVGLGYPNGNGGRTYLLDKSL--NVAYNESALSTQKATY 335
Db 405 TALLETADIYAPKEIADIAVGA-----GRVNYKAIKYDDYAKLFTGSVADKGSATH 457
Qy 336 TETATACKPLKISLVMSDAPASTASVTLVNDLVLITANGRTYVNDSSAFPDNNWDG 395
Db 458 TFDVSGATFVATLYMD-----TGSSDIDLVLVYDPNGE-----VDYSYTA 498

Qy 396 RNNVENVFNPSQGTITIEVCANVPEPONESLAIYN 434
Db 499 YGGEKVGYNPNFAGTWTYKVASK--GAANYQVDVVS 534

RESULT 12
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.8%; Score 423; DB 13; Length 412;
Best Local Similarity 29.6%; Pred. No. 1.5e-28;
Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

Qy 18 GLYGGQIVAVADTGLDTRNDSMEAFRGKITALYALGRTNAN-----DTNGHGH 71
Db 22 GYDSGIGITIGITIDID-----ASHPDLOGKV-----IGWPFVNRSYDYDHGHGH 70
Qy 72 VAGSVLGNATNK--GMAPOANLVFQSIM--DSSGGLGGLPSNLQTLFSAQFA 125
Db 71 VASIAAGTGAASNGKTYGMAPGAKIYVLGADGSSISTIKGVMAVANDKDYIKY 130
Qy 126 HTNSGA-----AVNGAYTTDSRNVDDYRKNDMTLLFAAGNEGNGGTISAP 173
Db 131 INLSGSSSDGTDLALQAVNNAMDA-----GLVYVVAAGNGGPNKYTTIGSP 178
Qy 174 GTAKNAITVATENLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVMAAGTYILSAR 233
Db 179 AAASKVITVGA-----VSKYDITFSRGPTADGRIRKPEVVAEGNWIIPARAS 227
Qy 234 LAPSSFWANHSKYAYMGSTVATPIVAGNVAQLREHFVNKGITPK--PSLLKALIA 291
Db 228 ---GTSWGPINDYTYAAGTSMATPHVAGIALLQ-----AHSWTPDKYKALIE 277
Qy 292 GA-----ADVGLGYPNGNGGRTYLLDKSLNVAYNESALSTQK-----TYTFT 338
Db 278 TADIVKPEIADIAVGA-----GRVNAVAIN--VDNRAKLFTGYVAKKSGQTQOFV 328
Qy 339 ATAGKPLKISLVMSDAPASTASVTLVNDLVLITANGRTYVNDSSAFPDNNWDGRNN 398
Db 339 ISGASFVATLYMDAN-----SDLDLYYDPNGN-----VDYSYTA 369
Qy 399 VENVFNPSQGTITIEVCANVPEPONESLAIYN 434
Db 370 FEKVGYNPNFAGTWTYKVAS--GSANYQVDVVS 402

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:


```

; APPLICANT: TAKAKURA, Hi-karu
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOUO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIORITY FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIORITY FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIORITY FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4

Query Match      18.8%; Score 423; DB 13; Length 522;
Best Local Similarity 29.6%; Pred. No. 2.1e-28;
Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

QY 18 GLYGQGIIVAVADTGLDTGRNDSMEAFRGKITLALYALGRINNAN-----DTNGHGH 71
DB 22 GYDSSGITTIGITDID-----ASHPDLOGKV-----IGWDFVNGRSYPPYDDHGHGH 70
QY 72 VAGSVLGNAGTN-----KGAPQANLVFQSIM--DSSGGLGGLPSNLQTLFSGAFSAGARI 125
DB 71 VASIAAGTGAASNGKYGAPGAKLAGIKVLGADGSGSISTIKGYEMAVDNKKDKYIKV 130
QY 126 HTSMWGA-----AVNGAYTTDSRNVDYVRKNDMTILFAGNEGPNGGTISAP 173
DB 131 INSLGSSQSSDGTDLAQVNAAMDA-----GLVVVAAGNSGPNKYITIGSP 178
QY 174 GTAKNALTIVGATENTLRPSFGSYADNINVAQFSSRGPTRGRKPDYVAPGTYILSARS 233
DB 179 AASKSVITVGA-----VDKVDVITSSFSRGPTRADGRKPEYVAPGMMITIAARS 227
QY 234 LAPDSSFWANHDSKYAVMGTSMTATPIVAGNVAQLREHFYKRGITPK--PSLLKALIA 291
DB 228 ---GTSMGQPIINDYTTAAPGTSMTATPIVAGIALLLQ-----AHPSWTPDKVKTALIE 277
QY 292 GA-----ADVGLGYPNGQMGRTYLDKSLNVAVNESSALSTSOKA-----TYTFT 338
DB 278 TADIVKDEIADIAYGA-----GRVNAKAIN--YDNVAKLVFTGYVANKGSQTHQFV 328
QY 339 ATAGKPKLISLVMSDAPASTTASVTLVNDLVLTPANGTRYVGNDFSAFPDNNMDGRNN 398
DB 329 ISGASFTATLYDNAN-----SDLDLYLPNGNQ-----VDYSYTAAYYG 369
QY 399 VENVFINSPOSQTYTIEVQAYNVVPGPNFSLAIYN 434
DB 370 FEKVGYYNPTDGTWTIKVSYSS---GSANYQVUVVS 402

RESULT 14
US-10-090-624-16
; Sequence 16, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hi-karu
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOUO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

```

```

; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIORITY FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIORITY FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIORITY FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match      18.8%; Score 423; DB 13; Length 654;
Best Local Similarity 29.6%; Pred. No. 2.9e-28;
Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

QY 18 GLYGQGIIVAVADTGLDTGRNDSMEAFRGKITLALYALGRINNAN-----DTNGHGH 71
DB 154 GYDSSGITTIGITDID-----ASHPDLOGKV-----IGWDFVNGRSYPPYDDHGHGH 202
QY 72 VAGSVLGNAGTN-----KGAPQANLVFQSIM--DSSGGLGGLPSNLQTLFSGAFSAGARI 125
DB 203 VASIAAGTGAASNGKYGAPGAKLAGIKVLGADGSGSISTIKGYEMAVDNKKDKYIKV 262
QY 126 HTSMWGA-----AVNGAYTTDSRNVDYVRKNDMTILFAGNEGPNGGTISAP 173
DB 263 INSLGSSQSSDGTDLAQVNAAMDA-----GLVVVAAGNSGPNKYITIGSP 310
QY 174 GTAKNALTIVGATENTLRPSFGSYADNINVAQFSSRGPTRGRKPDYVAPGTYILSARS 233
DB 311 AASKSVITVGA-----VDKVDVITSSFSRGPTRADGRKPEYVAPGMMITIAARS 339
QY 234 LAPDSSFWANHDSKYAVMGTSMTATPIVAGNVAQLREHFYKRGITPK--PSLLKALIA 291
DB 360 ---GTSMGQPIINDYTTAAPGTSMTATPIVAGIALLLQ-----AHPSWTPDKVKTALIE 409
QY 292 GA-----ADVGLGYPNGQMGRTYLDKSLNVAVNESSALSTSOKA-----TYTFT 338
DB 410 TADIVKDEIADIAYGA-----GRVNAKAIN--YDNVAKLVFTGYVANKGSQTHQFV 460
QY 339 ATAGKPKLISLVMSDAPASTTASVTLVNDLVLTPANGTRYVGNDFSAFPDNNMDGRNN 398
DB 461 ISGASFTATLYDNAN-----SDLDLYLPNGNQ-----VDYSYTAAYYG 501
QY 399 VENVFINSPOSQTYTIEVQAYNVVPGPNFSLAIYN 434
DB 502 FEKVGYYNPTDGTWTIKVSYSS---GSANYQVUVVS 534

RESULT 15
US-10-112-488-39
; Sequence 39, Application US/10112488
; Publication No. US20030082746A1
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yoshimi
; APPLICANT: UMEZAWA, Yukiko
; APPLICANT: YOKOYAMA, Keiichi
; APPLICANT: MATSUI, Hiroshi
; TITLE OF INVENTION: PROCESS FOR PRODUCING TRANSLUTAMINASE
; FILE REFERENCE: 219286USOCONT
; CURRENT APPLICATION NUMBER: US/10/112,488
; PRIORITY FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: PCT/JP00/06780
; PRIORITY FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: J2000-280098
; PRIORITY FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: JP11-280098
; PRIORITY FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 70

```

SOFTWARE: PatentIn version 3.1
SEQ ID NO 39
LENGTH: 1079
TYPE: PRT
ORGANISM: Streptomyces albobogisecius
US-10-112-488-39

Query Match 16.3%; Score 365; DB 14; Length 1079;
Best Local Similarity 31.3%; Pred. No. 7.7e-23;
Matches 146; Conservative 56; Mismatches 177; Indels 88; Gaps 19;

```
QY 3 VAR----GIYKADVAQS-----SYGLYGGQIVAVADTGLDTGRNDSMHEAFRG 48
Db 160 VARVWLDGVKAKASLDTSVGQIGTEKAWAGYDGGGVKIALVLDGV-----ATHRDLKG 213
QY 49 KITALLALGRNTNANDPENGHGHVAGSVLNGA---TNKMAPQANLVFQSIMDSSGGL 104
Db 214 QVTASKNFTSAPFTTGVDVGHGTHVASIAAGTGAQSKGYKGVAPGAKILNGKVLDDAG-- 271
QY 105 GGLPSNLQTLFSSQAFSAGARIHTNSWGAAVNGAVTTDSRYVDVYRK--NDMTILF--AA 160
Db 272 FGDDSGILAGMENAAGADIYVMSLG---GMDTPEPTDLEAAVDKLSAKGILFAIAA 327
QY 161 GNEPENGGISAPGTAKNAITVGATENTLPSFGSYADNINHVAFSSRGP-TYDGRIKPD 219
Db 328 GNEGPQ--SIGSPGSADSAALTGA-----VDDKDKLADFSTGTGRLDGDGAVKPD 374
QY 220 VMAPGTIILARSGLAPDSFFMANHDSKYAVMGTSNATPPIVAGNVVOLREHFVKRGIT 279
Db 375 LTAGVDTITAAAKANDIAKEVEKEPRAGYMTISGTSNATPHVAGAAALLKQCHPE----- 429
QY 280 PKESLLKALLIAGADYGLGYENGNGWGRVTLDKSLNAVAYNESSALS----- 328
Db 430 WKYAEELKAGALTASTKDQ--KYTPFEQSGSGRVQVDKAITQTVIAEPVLSFQVQWPHADD 487
QY 329 --TSOKATYTFETATAGKPLKISLWSD----APAS--TTASVTLVNDLDVITAP-NGT 378
Db 488 KPYTKULTYRNLTGEDVTLKLTSTATGPKGAAPAGFTLIGASTL-----TVPANGT 539
QY 379 RYVGNDPSAPFDNNMDGRNNVENVFINSPOS---GTYTIEVQANYV 421
Db 540 ASYDVVTADTRLGAGAVDGTYSAYVVAATGAGQSVRTAAVEREVESYV 586
```

Search completed: March 31, 2004, 16:33:58
Job time : 41.2177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3508 Seconds

(without alignments)
3677.911 Million cell updates/sec

Title: US-09-985-689A-2

Perfect score: 2245

Sequence: 1 NDVARGIVADVAQSSYGLY.....EYQAVNVPQPNLSLATVN 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	526.5	23.5	1743	2 T18279	multidrug resist
2	490	21.8	1905	2 T18267	intracellular alk
3	357.5	15.9	444	2 B83891	intracellular alk
4	326	14.5	442	2 A69587	microbial serine p
5	302.5	13.5	806	2 A41341	pyrolysin (EC 3.4.
6	302.5	13.5	1398	2 T28159	pyrolysin (EC 3.4.
7	289	12.6	419	1 S28535	subtilisin (EC 3.4
8	278	12.4	799	1 G83753	subtilisin-type pr
9	278	12.4	420	1 S23407	subtilisin (EC 3.4
10	277	12.3	580	2 S11890	serine proteinase
11	274.5	12.2	1345	2 T29090	surface layer-asso
12	270	12.0	715	2 JC4808	alkaline serine pr
13	260	11.6	1331	2 A72547	probable surface l
14	259.5	11.6	513	1 A35742	equalysin (EC 3.4.
15	259	11.5	627	1 D75393	serine proteinase
16	254.5	11.3	757	2 C84120	subtilisin-type pr
17	252.5	11.2	534	1 JS0173	alkaline proteinas
18	250.5	11.2	402	1 JU0332	cell wall-associat
19	249.5	11.1	894	2 F69730	alkaline proteinas
20	245	11.0	380	2 A49778	high-alkaline seri
21	245.5	10.9	401	2 T39874	serine proteinase
22	241	10.7	488	2 A11930	protease (import
23	240	10.6	378	2 A33973	high-alkaline seri
24	239	10.6	382	2 I39780	subtilisin (EC 3.4
25	238	10.6	519	2 S71451	halolysin R4 (EC 3
26	234.5	10.4	379	1 SUBSCL	subtilisin (EC 3.4
27	234.5	10.3	382	1 SUBSN	subtilisin (EC 3.4
28	232	10.3	321	1 S27501	alkaline proteinas
29	231.5	10.3	1052	2 T17093	intraluminal subti

30	230.5	10.3	613	2 S75576	hypothetical prote
31	229.5	10.2	381	2 JH0778	subtilisin (EC 3.4
32	229	10.2	1167	1 A35066	streptococcal C5a
33	228.5	10.2	525	2 G84406	halolysin (importe
34	228	10.2	1374	2 D72593	hypothetical prote
35	227.5	10.1	381	1 SUBS1	subtilisin (EC 3.4
36	226	10.1	384	1 UC4802	alkaline proteinas
37	225.5	10.0	272	2 A23624	subtilisin (EC 3.4
38	225.5	10.0	275	2 UC1085	subtilisin (EC 3.4
39	225.5	10.0	381	1 SUBS	subtilisin (EC 3.4
40	225.5	10.0	381	2 JQ1487	subtilisin (EC 3.4
41	225.5	10.0	910	2 C69456	subtilisin sendai
42	224	10.0	535	2 B82358	alkaline serine pr
43	222.5	9.9	769	2 D86335	T20H2 6 protein
44	219.5	9.8	274	1 SUBS2	subtilisin (EC 3.4
45	219.5	9.8	361	2 A48373	high-alkaline seri

ALIGNMENTS

RESULT 1	18279	multidrug resistance transport protein - slime mold (Dictyostelium discoideum)
C:Species:	Dictyostelium discoideum	
C:Date:	15-Oct-1999	#sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession:	T18279	
R:Shaulsky, G.; Loomis, W.F.		
Submitted to the EMBL Data Library, June 1996		
A:Reference number:	Z18855	
A:Accession:	T18279	
A:Status:	Preliminary; translated from GB/EMBL/DBJ	
A:Molecule type:	mRNA	
A:Residues:	1-1743 <SHA>	
A:Cross-references:	EMBL:U60086; NID:G1399214; PID:G1399215; PIDN:AA03331.1	
C:Genetics:		
A:Gene:	tagC	
Query Match	23.5%; Score 526.5; DB 2; Length 1743;	
Best Local Similarity	27.8%; Pred. No. 9.8e-26;	
Matches	164; Conservative 82; Mismatches 152; Indels 191; Gaps 22;	
QY	19 LYGGQIVAVADTGLDGR---NDS-----SMHEFRKRIALVALGRTNNANTNGH 68	
DB	314 LRGGQILSLADYGLDSHCFFSDSKYPIPLNSVNLNHR-KVATYITTSDDSDKVDGH 372	
QY	69 GTHVAGSVLG-----NGATNKGMAPQANLVQSIWDSGGGLPSNLQTFQSAFSA 121	
DB	373 GTHICGSAAGTPEDSSVNISSFGLATDAKIAFPDLASGSSSLTP-PSDLKOLYPLIDA 431	
QY	122 GARHTNSWGA---AVNGAYTTDSRVNDYVRKN-DMTLLFPAAGNBPNGGTIS-AFG 174	
DB	432 GARVHCDMSVGSVEGYGYSPTASIDPLFTHPDIILRAAGN---NEQYLSILTQS 488	
QY	175 TAKNAITVGAATENLR-----PSFGSYADNI----- 199	
DB	489 TAKVITVGAHQTHENYVLIDGRVNTYQSSVDINQSLIDPDSRYCNVTTAOCCEESNA 548	
QY	200 -----NHVAQFSSRGFTKQGRIRPDVMAQETYLISA 230	
DB	549 TTGLASCCPTLLRKSVIDAANTOPLLYNENNISFSSKGPETHGRMPALVAGEYITSA 608	
QY	231 RSSLA-----PDSSFWANDSKAYVGTSMATPIVAGVAQUREH-----FVK 274	
DB	609 RSNGANITTDGCGDSL-PNINNALIA-ISGISMATSPFAAATTLIRQYLVGVGYPTGSIVE 666	
QY	275 NRGITPKRSILKAALIGA-----ADVGLGYNGN-----QGWRYTLD 313	
DB	667 SNKQPGSLIKALMINNAOLINGFOLITSSITVPSNVGFENFAGASLVQGWGALRMS 726	
QY	314 KSLVAVYVNESS-----ALTSQAKATYET----- 338	
DB	727 NWLHVAVNNNSNNNNKTSQGITVFDGIGGLDLVKNQMKESLSIGQNTSICFTYKPS 786	

QY 339 ---ATACKPLK---ISLWSDAPASTTASVTLVNDLDTLTAENGRTYVGNDFSAFDDNN 392
 Db 787 SSSSSGNINPRVAVATLWTDPSYAGAKFNLVNNDLTLMI-----YKDNOSTIFYNN 840
 QY 393 WDGR-----NNVENVFINSPOSGTYTIEVQAVNVPVGPONFS 429
 Db 841 QGSSSFLGLAPDTDTLNNBGIYVHPTEPWTIFPMVAGINVPVGPONFS 889

RESULT 2

multidrug resistance protein - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18267
 R:Shaulsky, G.; Kuspa, A.; Loomis, W.F.
 Submitted to the EMBL Data Library, January 1995
 A:Description: An MDR transporter/serine protease gene is required for prestalk speciali
 A:Reference number: Z18850
 A:Accession: T18267
 A:Status: preliminary; translated from GS/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-1905 <SHA>
 A:Cross-references: EMBL:U20432; NID:9664839; PID:9664840; PIDN:AAA62212.1
 C:Genetics:
 A:Gene: tagB

Query Match 21.8%; Score 490; DB 2; Length 1905;
 Best Local Similarity 28.0%; Pred. No. 2,4e-23;
 Matches 162; Conservative 74; Mismatches 159; Indels 184; Gaps 23;

QY 19 LYGGQIVAVADTGLDTCR---NDS-----SMHEAFRGKITALYALGRTNANDTNGH 69
 Db 376 LRGGQILSTADTGLDSSHCFPSDKYPIFNQVNEHRAVIV---YITHDHEDEVNH 432
 QY 69 GTHVAGSVLNG-----ATNKMAPQALVQSINDSGGIGL-PSYLQTLFEOQAS 120
 Db 433 GTHVCGSAAGPEDSSVAISSFGLATDAKIAFDL---SSGSSEPTPEDEYQMYPLVD 490
 QY 121 AGARHTNSWGA---AVNGAYTTDSHNDVYRK-NDMTILFAAGNEGNGGTTISAPGT 175
 Db 491 AGARVHDSDWGSVELQGYGYGSDDAGIDAFLEYEFESILLAAQN-NELFRSLAAQNT 549
 QY 176 AKRAITGATENLFPSTGS-----YADNI----- 199
 Db 550 AKRAITGAEQTAHVNVSDALEYDPSDNANFQPCLPDKKYCNVTTAKCSEVSNVKG 609
 QY 200 -----NHVAFSSRGPTKDGRIKPDVMAFGTYILSARS 233
 Db 610 LQCCPASIKQNASDFTTQPOPYNNENMGSSFGKTHGRILKPDIVAGEYITTSARSN 669
 QY 234 -----IAPSSFWANHDSKYAVWGTSMAPIVAGNVALRHF-----YKNG 277
 Db 670 GENSTDOCQGGSTL--PVANGILMSISGTSMAPIPLATAATTLRQYLVGYPTGESYEENK 727
 QY 278 ITPKPSLLKALILGADVGLGY-----PNGNGMGRTVLDSLVN 318
 Db 728 LIPFGSLKALMINNQLNGTYWMSASTNPNAITEQINGANLQGMALMMN---NW 784
 QY 319 AYVNESS-----ALTSOKAT-----YFFT-----ATAG 342
 Db 785 LYVKSNSPTPSRWIGIGLGKNGKATEWKEDLSLGSINKSYCFITYPSSSSSGSGGGG 844
 QY 343 KP-LKISLVMSDAPASTTASVTLVNDL-----VITAN--GTRYVGNDFSAFDD 390
 Db 845 TPRVATLWTDPSYSGAKFNLVNNDLILNSDDSIITLIGSGSLOPAKVAQP--- 902
 QY 391 NMDGRNVNVPVINSPOSGTYTIEVQAVNVPVGPONFS 429
 Db 903 ---DLNNVEGIIINPTKANMYKFTAGTVNVPISQPKFS 938

RESULT 3

B83891
 Intracellular alkaline serine proteinase aprx [imported] - Bacillus halodurans (strain
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: B83891
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: B83891
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-444 <STO>
 A:Cross-references: GB:AP001513; GB:BA000004; NID:910174345; PIDN:BA05649.1; GSPDB:GNO
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: aprx

Query Match 15.9%; Score 357.5; DB 2; Length 444;
 Best Local Similarity 30.2%; Pred. No. 1.2e-15;
 Matches 110; Conservative 52; Mismatches 105; Indels 97; Gaps 16;

QY 11 DVAQSSVGLYGOQIVAVADTGLDTCGRNDSMEHAFRGKITALLY-ALGRTNANDTNGH 69
 Db 135 EVIRNGERTLTGKDVTAIVDTGI-----YFHEDLEGRIKAFVDVFNQREBPYDDGNG 187
 QY 70 THVAGSVLNGATN---KMAPQALVQSINDSGGIGLPSNLO----- 112
 Db 188 THCGADAAAGASSDGGQYRGPAPFANVIGYVLKQ-QMGSLSEIMQGVEMCIQYNEHP 246
 QY 113 -----TLFSQAF-----SAGARHTNSWGAAVNGAYTTDSHNDVYRKNDMTI 156
 Db 247 DDEHILISMISLGGQALPVENQEDPMVRIVEBANNAIG-----TV 286
 QY 157 LFAAGNEGNGGTTISAPGTAGTATTTGATENTLPSFSYADNINHVAGQSSRGPTKGR 216
 Db 287 CVAAGNSGPDQQTATLSPVSEKVTIVGALDD-RITTRRDD---VAFSSRGPTIYKGP 342
 QY 217 KPDVAPGTIYLSARSSILAPDSF-----WANHDSKYAVWGTSMAPIVAGNVALRHF 271
 Db 343 KPDILAPGVNIVSLRS---PNSFYDKIQKSGRVSHTTMSGTSMAPIVAGNVALRHF 399
 QY 272 FVKRNGITPKPSLLKALILGADVGLGYENGNGMGRTVLDSLVN---AYVNESSALS 328
 Db 400 ---EPNLTPDE--VKTRLM-----EGTDRWA---DRDPNYGAGYISAGCAIP 439
 QY 329 TSOK 332
 Db 440 NSEE 443

RESULT 4

A69587
 Intracellular alkaline serine proteinase aprx - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: A69587
 R:Kumet, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
 C:Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A:Erlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 ich, J.; Harwood, C.R.; Hentz, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapides, A.; Lardinois
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.W.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y, M.; Ogawa, K.; Ogawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
 A:Authors: Schleich, S.; Schroeter, R.; Seckforn, F.; Sekiguchi, J.; Sekowska, A.; Sero
 akuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetta, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
 A:Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Zundstein, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: A69587
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-442 <KUN>
 A:Cross-references: GB:299113; GB:AL009126; NID:g2634090; PIDN:CAB13610.1; PID:e1183385;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: aprX
 C:Superfamily: subtilisin homology
 F:146-398/Domain: subtilisin homology <SBT>
 Query Match 14.5%; Score 326; DB 2; Length 442;
 Best Local Similarity 30.6%; Pred. No. 1,2e-13;
 Matches 103; Conservative 45; Mismatches 97; Indels 92; Gaps 14;
 QY 11 DVAGSSYGLGGGQGVAVADTGLDGRNDSNHEAFRGKITALVALGRTNNAN-----D 64
 Db 136 EVVRNGQTLTGKGVAVAVDTGI-----YHPDLEIRI-----IGPADVWQKTEPYD 183
 QY 65 TNGHGTAVAGSVLGGGATN-----KGMAPQANLVFQSIWDSGGGLGLPSNLTQLFSQAFS 120
 Db 164 DKGHGTGAGDVAVSSGASGGYRGPAPBANLIGVKVANKGS-CTLADILEGV----- 236
 QY 121 AGARLHTNSWGAAN-----GAYTTSRNVVDYRKND-----MTIL 157
 Db 237 -----EMCIQVNEEDNEPEIDIMSMSLGGDALRYHEGDPVRAVEASAGIYVC 288
 QY 156 PAAGNEGPGGTISAPGTAKNAITVGAENTLPSFGSVADNNHVAQSSSGPTKDGRIK 217
 Db 289 VAAAGNSGPGSQTIABPVSEKVIITGALDNNMTA-----SSDDTIVASSSSSGPTIYGEX 344
 QY 218 PVWAPGTYILASRSLAPDSF-----WANHDSRYAVMGTSWATPIVAGNVAQREHF 272
 Db 345 PDLAPGVNIIISLR---PNSYIDKLKSSRVSQSYFTWSGTSMATPICAGIALILID-- 399
 QY 273 VKNRGITPK--PSLKAALINAGADVGLCPNGNGM 307
 Db 400 -QNPDLTPDEVKELK-----NGTTKW 420
 RESULT 5
 A41341
 Microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 17-Jul-1992 #sequence, revision 17-Jul-1992 #text_change 20-Jun-2000
 C:Accession: A41341; B41341; S39700; D69730
 R:Sioma, A.; Ruto Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.
 J. Bacteriol. 173, 6889-6895, 1991
 A:Title: Cloning and characterization of the gene for an additional extracellular serine
 A:Reference number: A41341; MUID:92041574; PMID:1938892
 A:Accession: A41341
 A:Molecule type: DNA
 A:Residues: 1-806 <SLO>
 A:Cross-references: GB:M76590; NID:G143819; PIDN:AAA22881.1; PID:G143820
 A:Accession: B41341
 A:Molecule type: protein
 A:Residues: 161-195 <SL2>
 R:Glaeser, P.; Kunst, F.; Anand, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
 A.; Rapoport, G.; Darchin, A.
 Mol. Microbiol. 10, 371-384, 1993
 A:Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
 A:Reference number: S39655; MUID:95020537; PMID:7934828
 A:Accession: S39700
 A:Molecule type: DNA
 A:Residues: 1-806 <GLA>
 A:Cross-references: EMBL:X73124; NID:G413923; PIDN:CAA51601.1; PID:G580871
 R:Kunert, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alioni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruchet, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Enllich, S.D.; Emmerson, P.T.; Entian, K.D.; Erington, J.; Fabret, C.; Ferrati, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 tech, J.; Harwood, C.R.; Hanaul, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koertter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
 A:Authors: Schleicher, S.; Schreier, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Serot
 Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, i
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Darchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: D69730
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-806 <KUN>
 A:Cross-references: GB:299123; GB:AL009126; NID:g2634240; PIDN:CAB15835.1; PID:g26344
 A:Experimental source: strain 168
 C:Comment: The amino terminal sequence of the mature protein and a molecular weight of
 C:Genetics:
 A:Gene: vpr
 A:Start codon: TTG
 C:Superfamily: microbial serine proteinase vpr, subtilisin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-28/Domain: signal sequence #status predicted <Sig>
 F:29-160/Domain: propeptide #status predicted <Pro>
 F:160-348/Domain: subtilisin homology #status atypical <SBT>
 Query Match 13.5%; Score 302.5; DB 2; Length 806;
 Best Local Similarity 23.2%; Pred. No. 8.7e-12;
 Matches 127; Conservative 58; Mismatches 161; Indels 201; Gaps 17;
 QY 18 GLYGQGVAVADTGLDTR-----NDSNHEAFRGKITALVALGRIN 60
 Db 177 GYTGKGIKVAIDTGVVNHPPDIKKNFGQYGVDFVNDYDPERPTG-----D 225
 QY 61 NAMDNGHTGTHVAGSVLGGATNKGAPQANLVFQSIWDSGGGLGLPSNLTQLFSQAFS 120
 Db 226 PREGADHGTAVAGTAAAG-TIKGVAPOKTLAVVLGPGG--SGTENVAVGVRVAVQ 282
 QY 121 AGARLHTNSWGAANVATYDTSKRVDDYRKNDMTILPAAGNEGPGGTISAPGTAKNAI 180
 Db 283 DGADVNNLSLGNLNPDMATSTAL-DWMSSEGVAVTNSNGSGPMTVSGPSISREAI 341
 QY 181 TVGATE-----NLSPFGSY----- 195
 Db 342 SVGATQPLNEVAVTGTSGSYSAKVMGYNKEDDYKALNKEVELVEAGIGBAKDFEKKLT 401
 QY 196 -----ADNI----- 199
 Db 402 GKVAVVRGSIAPVDKADNKKAGAGIMVVYNNLSGEIENAVPMGMSVPTIKLSLEDGEXL 461
 QY 200 -----NHVAFSRRGPTK-GRIKPVWAPGTYILASRSLAP 236
 Db 462 VSALKAGETTTTKLTVSKALGQVADFSRRGVNMTMIKPIASPGVNIIVTITTHDP 521
 QY 237 DSSEFWANHDSRYAVMGTSWATPIVAGNVAQLREHFVKNRGITPKPSL-LKAALTAGAA 294
 Db 522 D-----HPYGSQSKQSTWASPHIAGAVALIKO-----AKPKMSVEQIKALINMTAV 568
 QY 295 DV-----GLGVPNGNGGRTLLDKSLNVAVNNSSALSTQKATYFTATAGAPLISLV 350
 Db 569 TLKSDSEVYPHNAQAGSARI--NNA--IKADSLVSPSISYGFILKNGMETKNETF 623
 QY 351 WSDAPASTASVTLVNDLVLITAPNGTRVYVNGDFAPEDNNMDGNNVENVFINSPOG 410
 Db 624 TIENQSSIRKSYTL-----EVS--FNGSGISTSGTSRAVVIPIAHQTG 662
 QY 411 TYTLEVG 417
 Db 663 KATAKVK 669
 RESULT 6
 T28159

Pyrolysin (EC 3.4.21.62) - *Pyrococcus furiosus*

C:Species: *Pyrococcus furiosus*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000

C/Accession: 128159

R/Voorwerf, W.G.B., Eggen, R.I.L., Geerling, A.C.M., Platteau, C., Sieren, R.J., Vos, J. Biol. Chem. 271, 20426-20431, 1996

A/Title: Isolation and characterization of the hyperthermostable serine protease, pyrolysin

A/Reference number: 220481, MUID:96355370, PMID:8702780

A/Accession: 128159

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1398 <VO>

A/Cross-references: EMBL:U55835, NID:G1556462, PID:G1556463, PIDN:AA09761.1

A/Experimental source: DSM3638

A/Genetics:

A/Keywords: hydrolase; serine proteinase

Query Match 13.5%; Score 302.5; DB 2; Length 1398;
Best Local Similarity 26.3%; Pred. No. 1.8e-11;
Matches 138; Conservative 57; Mismatches 177; Indels 153; Gaps 22;

```

QY 21 GGGIVAVADTGLDTRGRDSSM-EMHAFRGKITALYALGRTNNAN-----63
DB 301 GNGVDLYVDITDIDYDTDEVPAGQVNTVDVAVFSTYGLPLVLAIEDPGEIYAVFGW 360
QY 64 DTNGGTHVAGVYLGNGATN-----KGMAPQAN 91
DB 361 DGHGHTHVAGVAGVYDSDNDAMDWLSMYSGEWEVESRLYGMQDYNTVITDVGVAFGAQ 420
QY 92 LVPGSINDSGGGLGSPNSIQTLFSAFSGARITHNSMAANGAYT--TDSRN--VVD 147
DB 421 IMAIRVL-RSDGRSMWDITLGM-TYATHGADVIVSKLGG--NADYLDGTPESVAVDE 476
QY 148 YVRKNDITLFAAGNEBPNNGTISAPGTAKAATVGTATNLRPSFGSYAD-----197
DB 477 LTKYGVAVFLAAGNEBPNNGTISAPGTAKAATVGTATNLRPSFGSYAD-----197
QY 198 -----NINHYAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSFFMANHDSKYA 249
DB 536 FYVFPAYTNV-RIAFSSRGPRIDGELKPVVAVAGVYSSLPWMTGADF-----585
QY 250 YMGSTMATPIVAGNVAYQLREHFVKNKGIPTKPSLLKAAIAGAA-----DVGIGYPNG 303
DB 586 -MGSTMATPIHVSQVALLISG-PKPEGIYVNDIILKYLESGATVLEGGPYGQXYTEL 643
QY 304 NQCKGRVTLDKSLNVAVNESSALSTSQKATYFTTAGKPLKISLVSDAPASTTA---360
DB 644 DQGGVAVNVTKWEI-----LKAINGTTLPIVDHMADKSYSDPAEYL 685
QY 361 SVTLVNDLDTVITAPN---GTRYVGNDFGAPFD---NNW-----DG-----RNNVE---400
DB 686 GVDVIRLVANRNPDIYEMHIKYVGTERTREIYATEWIKPIFVSGVILENTEFVL 745
QY 401 -----NVFINSPOSGTYTIEVQAVNVPGPQNF 429
DB 746 RVKXVDEGLEPGLVYGRIIIDPFL-TPVIEDEILNTIVTPEKFT 788

```

RESULT 7

S25835

subtilisin (EC 3.4.21.62) precursor - *Bacillus* sp. (strain TA41)

C/Species: *Bacillus* sp.

C/Date: 22-Nov-1993 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999

C/Accession: S25835

R/Davall, S.; Feller, G.; Narinx, E.; Gerday, C.

Gene 119, 143-144, 1992

A/Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph *Bacillus*

A/Reference number: S25835, MUID:93012986, PMID:1139802

A/Accession: S25835

A/Molecule type: DNA

A/Residues: 1-419 <DV>

A/Cross-references: EMBL:X63533, NID:G40198, PIDN:CAA45096.1, PID:G40199

C:Superfamily: subtilisin; subtilisin homology

C/Keywords: extracellular protein; hydrolase; serine proteinase

F11-23/Domain: signal sequence #status predicted <SIG>

F124-110/Domain: propeptide #status predicted <PRO>

F111-419/Product: microbial serine proteinase #status predicted <MAT>

F135-373/Domain: subtilisin homology <SBT>

F144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 12.6%; Score 283; DB 1; Length 419;
Best Local Similarity 33.0%; Pred. No. 6.6e-11;
Matches 87; Conservative 37; Mismatches 98; Indels 42; Gaps 12;

```

QY 21 GGGIVAVADTGLDTRGRDSSM-EAFRGKITALYALGRT--NNADTNGHTHVAGSV 76
DB 135 GAGINIAVLDGVNTNHPDLSNNVEQCKD-----FTVGTFTNDSCTDRGQGHVAGSA 189
QY 77 LGNGATNK--GMAPQANLVFQSIM--DSSGGLGSPNSIQTLFSAFSGARITHN-SW 130
DB 190 LANGTSGVYVAPADLVAAYKVLGDDSGYADDAIEAIRHAGDQATLNTVAVIMSL 249
QY 131 GAAYVCAVTTDSRVVDYVERKNDITLFAAGNEBPNNGTISAPGTAKAATVGTATNLRP 190
DB 250 GSSGESSLIT--NAVYAVDKVGLIAANGNSPKRGSIGYGALVNVAVALLENTIQ 306
QY 191 SFGSYADNINHYAQFSSRGPTKDG-----RIKPDVMAFGTYILSARSLAPDSSFFMANH 244
DB 307 N-GTY-----RVAPFSSRGHRRTAGDYVIQGDVEISAPAAVYST-----W-F 348
QY 245 DSKYAVMGSTMATPIVAGNVAYQL 268
DB 349 DGYATITSGTSMASPHAAGLAKI 372

```

RESULT 8

G83753

subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - *Bacillus halodurans*

C/Species: *Bacillus halodurans*

C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C/Accession: G83753

R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fuji, F.; Hir, Nucleic Acids Res. 28, 4317-4331, 2000

A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A/Reference number: A83650, MUID:20512582, PMID:11058132

A/Accession: G83753

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-799 <STO>

A/Cross-references: GB:AP001510, GB:BA000004, NID:G10173440, PIDN:BA04550.1, GSPDB:GNO

A/Experimental source: strain C-125

C/Genetics:

A/Keywords: hydrolase; serine proteinase

F11-29/Domain: signal sequence #status predicted <SIG>

```

QY 18 GLVGGQIVAVADTGLDTRGRDSSMHEAFRGKITALYALGRTNNANDT-----NGH 68
DB 171 GYTGSGITVALDVGVDTPHPD--LVHAF-GDYKGMDFIDNNDPQRTPEGDPGRTTTH 227
QY 69 GTHVAGSVLNGATNKGAPOANLVFQSIMDSGGLGSPNSIQTL--FSAFSGARITH 126
DB 228 GTHVAGTVANGLL-KGVAPDANILLAYRL-----GPGRGSTAGVIGIRAVQDGADIM 282
QY 127 TNSMAANGAVTTDSRVVDYVERKNDITLFAAGNEBPNNGTISAPGTAKAATVGTATE 186
DB 283 NISLGNLTINDPDPATSTAL--DMAAEGVAVTISNGNSGPNNTVGSPTSRDALSVAT- 340
QY 187 NLKPSFGSY-----ADNINH-----201

```

Query Match 12.4%; Score 279; DB 2; Length 799;
Best Local Similarity 25.0%; Pred. No. 2.8e-10;
Matches 119; Conservative 47; Mismatches 134; Indels 176; Gaps 17;

Db 341 --RLPNKYKAVFTSDGIDYPSADJINGPFSDELLTLDGETYEYAFAGLKPDEGEVD 398
 QY 202 ----- 201
 Db 399 VEGKIALIVAGEIPFEVKAENAKAAGAVGAILYNNVAGVPTVGLAIPITMISNEDGLK 458
 QY 202 -----VAOFSSRGPT-KDGIKPDVAPGTYILSARS5LA 235
 Db 459 MRNELENGQNTVTFSEIPDXLVGETVADPFSSRGVMTWIKIPDVAPGVAIVSTIPTHQ 518
 QY 236 PDSSFMANHDSKYAVMGSTMAPIPVAGNAQLREHVNRKRGITPRKSLIKALLINGAD 295
 Db 519 PDDPY-----GYSRGCTISVASHVAGAAALLIEAH-PWAGV---DHVKALAMNTAEN 567
 QY 296 V---GLGYPNGQGWGRVTLDRSLNVAAYNESSALSTQKATY-TFTATAGKPLX 346
 Db 568 LVDENGRMRYPHNTQAG-----SIRIVDAIESBTLVTPGSHSGFTTKERGRQVE 617

RESULT 9

S23407
 Subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
 C/Species: Bacillus sp.
 C/Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Jun-1999
 C/Accession: S23407
 R/Natrix, B.; Davail, S.; Feller, G.; Gerday, C.
 Biochim. Biophys. Acta 1131, 111-113, 1992
 A/Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarcctic
 A/Reference number: S23407; MUID:92256481; PMID:1581352
 A/Accession: S23407
 A/Molecule type: DNA
 A/Residues: 1-420 <NMR>
 A/Cross-references: EMBL:X62363; NID:940200; PIDN:CAA44227.1; PID:940201
 C/Genetics:
 A/Gene: studi
 C/Superfamily: subtilisin; subtilisin homology
 C/Keywords: extracellular protein; hydrolase; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-111/Domain: propeptide #status predicted <PRO>
 F:112-420/Product: microbial serine proteinase #status predicted <MAT>
 F:116-374/Domain: subtilisin homology <SRT>
 F:145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 12.4%; Score 278; DB 1; Length 420;
 Best Local Similarity 33.2%; Pred. No. 1.4e-10;
 Matches 88; Conservative 36; Mismatches 95; Indels 44; Gaps 13;

QY 21 GGGQIVAVADTGLDTGRNDSMEAFRGKITAL--VALGRT--NNANDTNGHTVAGS 75
 Db 136 GGGGINIVLDTGVNTN-----HDLNNVEQCKDFVGTYYNNSCTBQGHGVAGS 189
 QY 76 VLNGATNKK--GNAPQNLVFGSIM--DSSGGLGCLPSNLQTLFSGAFSAGARIHTN-S 129
 Db 190 ALADGGGNGVYGVAPADLMAYKVLGDGSGVADDAIRARAGQATLNTKVTNMS 249
 QY 130 WGAIVNAGVYTTDSKXNDVYKXNDMTLLFAAGBENGNGTISAPGTAKAIVTGATENR 189
 Db 250 LGSSGESSLITNAVN--YSYNKGVLIIAAGNSGPQSGIGYPGALVNVAAVALEN-K 305
 QY 190 PFSFGSYADINHVAFSSRGPT-KDG-----RIKPDVMAAGTYILSARSLADSSFMAN 243
 Db 306 VENTTY-----RVADFSRGVSWTDGYAIOKGDVEISAPGAALYST-----W-- 348
 QY 244 HDSKIVAMGTSVATPIVAGNAQL 268
 Db 349 FDGGYATISGTSMSAPHAAGLAAKI 373

RESULT 10

S11890
 Serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. ca
 N/Alternate names: subtilisin-related proteinase
 C/Species: Xanthomonas campestris pv. campestris

C/Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 03-Dec-1999
 C/Accession: S11890
 R/Liu, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
 Mol. Gen. Genet. 220, 433-440, 1990
 A/Title: A multipurpose broad host range cloning vector and its use to characterise an
 A/Reference number: S11890; MUID:90251253; PMID:2187155
 A/Accession: S11890
 A/Molecule type: DNA
 A/Residues: 1-580 <Liu>
 A/Cross-references: EMBL:X51635; NID:948533; PIDN:CAA35962.1; PID:948534
 A/Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-AL
 C/Superfamily: subtilisin; subtilisin homology
 C/Keywords: extracellular protein; hydrolase; serine proteinase
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:168-423/Domain: subtilisin homology <SRT>

Query Match 12.3%; Score 277; DB 2; Length 580;
 Best Local Similarity 26.6%; Pred. No. 2.5e-10;
 Matches 126; Conservative 53; Mismatches 150; Indels 144; Gaps 22;

QY 21 GGGQIVAVADTGL---DTGRNDSMEAFRGKITALVALGRTNNAND----- 64
 Db 166 GGGIVAVIDTIGTSHDLNNLILAGDFISDATTAADGNDRSDNADGDMVAANCGA 227
 QY 65 -----TNGGTHVAGSVLNGCATNKGAPOA-----NLVFQSTMS 100
 Db 228 GIPAASSWHGTHVAGVAAVNTNTTGVACVAGKVPVAVLCKGGSLSIDADIYWA 287
 QY 101 SGG-UGCLPSNLQ--TLPSQAFSAGARIHTNMGAAVAGVATTSRVVDVYRKNDMTIL 157
 Db 288 SGGTYSGLPAAANPREVYNMISLGGGSGSTTMQN-ALNGAVSRRT-----TVV 334
 QY 156 FAAGNEGNGGTTISAPGTAKAIVGATEN--LPFSFGSYADINHVAFSSRGPTDGR 215
 Db 335 VAAQNDASNVGQ-SIPANCAVIVAAATTSAGAAVSXNFGTGL----- 377
 QY 216 IKPDVMAAGTYILSARS--LAPSSFMANDSXYAVMGSTMAPIPVAGNAQLREHFV 273
 Db 378 ---DVSABGSSILSTLNSGTTTPQS-----ASYASNGISVAPVAGVVALVQS--V 425
 QY 274 KNRGTPPK--PSLKK--AALTAGADVGLG-----PENGQGWGRVTLDK 314
 Db 426 APTALTTPAAVEITLKTARALPGACSGCGGIVNADAATAINAGSGGCGGCGNTLTN 485
 QY 315 SLNVAAYNESSALSTQKATYFTATAGKPKIKSLVMSDAPASTASVTL--VNDIDIV 371
 Db 466 GTPVT---GLGATGALNNTITVPAG-----SGTLVTTISGSGGDADLY 527
 QY 372 ITAPNGTRYVNDPSAPFDNNMDR---NNVENVFINSPOSGTYTTEVQAYN 420
 Db 528 VRAG-----SAPTDSAYTCRPYRSGNAETCTITAP-SGYVYRLKAYS 569

RESULT 11

T29090
 Surface layer-associated STABLE proteinase - Staphylothermus marinus
 N/Alternate names: hyperthermostable proteinase
 C/Species: Staphylothermus marinus
 C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C/Accession: T29090
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1345 <MAY>
 A/Cross-references: EMBL:U57968; NID:91374755; PID:91374756; PIDN:AA03223.1
 A/Experimental source: strain FL
 C/Function:
 A/Description: probably serves an exodigestive function related to the organism's energy

A>Note: stoichiometric S-layer component

Query Match 12.2%; Score 274.5; DB 2; Length 1345;

Best Local Similarity 29.3%; Pred. No. 1.1e-09;

Matches 97; Conservative 48; Mismatches 105; Indels 81; Gaps 14;

46 FRKITALYLGRTNNANDTNGHCTHA-----GSVL-----GNGKTNK--GMAQANLV 93

445 YQGRYIAL-----VSDFHGISTAVIASRGVLDYLDGKLYRIMGVAAPAKI- 495

94 FQSIMDSGGLGGLPSNLTQFQAFSAG-----ARHTNSM 130

496 -----AGDAMLNLGILVL--EAWLAGENITEBEDGVYSLDPFPHADIISNM 546

131 GAANVAGYTTDSKND-----DYRKNDMTLLFAAGNENGGITISAPGAK 177

547 GSIIYINFWLQFPIDYRSFMDLIRNYLLGDHVTIVFAAGNGPQYSNGAPGTGL 606

178 NATTVGATE--NRPSPG--SYADNINVAQPSRSGPTKGRIRKPDVAPGTIISARS 232

607 LVITAGASTLMDYTRIVYEGYAD--EYIPSSSGPTGGGPKPDIVNIGAFEMASTR 663

233 SLAPDSEFWANHDSKXAVMGSTMAPTVAGNVAQLREHFVKNRGITPKPSLKALLAG 292

664 TI-DGRGYAQPD-----VFQGTSEATPYTSGTLALVFOAKVEYVNTTDPVTAKIILKSS 718

293 AADVGLGYPNGNQMGKRVTLDKSLNVAVYNE 323

719 AKDI--WYPAFSGQSGRVDAKADIVFISE 747

RESULT 12

JC4908

alkaline serine protease (EC 3.4.-.-) I precursor - Alteromonas sp.

N:Alternate names: subtilase

C:Species: Alteromonas sp.

C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999

C/Accession: J04908

R:Unijob, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y.

BioSci. Biotechnol. Biochem. 60, 1284-1288, 1996

A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacter

A:Reference number: J04908; MUID:97141200; PMID:8987544

A:Accession: J04908

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-715 <TSU>

A:Cross-references: DDBJ:D38600; NID:91536787; PIDN:BAA18912.1; PID:d1019647; PID:g21602

C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sens

C:Genetics:

A:Gene: aprI

C:Superfamily: subtilisin homology

C:Keywords: hydrolase

F:1-40/Domain: signal sequence #status predicted <SIG>

F:41-150/Domain: amino-terminal propeptide #status predicted <ATP>

F:151-496/Product: alkaline serine protease I #status predicted <MAT>

F:182-452/Domain: subtilisin homology <SBT>

F:487-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 12.0%; Score 270; DB 2; Length 715;

Best Local Similarity 26.3%; Pred. No. 9.1e-10;

Matches 126; Conservative 49; Mismatches 168; Indels 136; Gaps 23;

21 GGGGIVAVADTG---LDGRNDSMHEAFRGKITLXALGRTNMND-----TNG----- 67

Db 182 GGGGIVAVADTG---LDGRNDSMHEAFRGKITLXALGRTNMND-----TNG----- 67

QY 68 -----HGTIVAG---SVLNGATNKMAPOANLVQSIIMDSGGGLGHP 108

Db 242 DSGGQPVPRADQDSWGTGTVAGTVAVNTNNGGVAGVADAKVVEVRVL---GKGGGLT 298

QY 109 SNIQTLSQAFSAGARHTHSWGAAY-----NGAYTTDSRVNDVYRKNDMTILFAAG 161

Db 299 SDIADGIIMVSGSDRVPANVAVNTNMSIGGGAGCSATQAINQARNNGTVIYAG 358

QY 162 NEGNNGGITSAPGAKNAITVGT--ENLPSGVSADNINVAQSSSRGPTKGRIRKPD 219

Db 359 NDNDNSANVN--PGNCQNVAVVASVGRGSGRAYSNGANI-----D 398

QY 220 VMAPGTIILARSGLAPDSFWANHDS-----KYAVGGSMAPTVAGNVAQLR-- 269

Db 399 VMAPG-----GAQSFADDPESGILSTHNSGSGAPANDSYHSQGTSMAPFAVAGVAAALIKQA 454

QY 270 -----EHFVKN--RGITPKPSLKALLAGADV--GLD-----YPNGNQMGKRVTL 312

Db 455 KPSATPDEVEITLKNTRSPAGSCNCGTGVVAAAANVADLDVYTPPTGN-----TL 508

QY 313 DKSINVAVNESSALSTQKATYFTATACKPKLISLWSDAPASTASVTL--VNDLD 369

Db 509 ED--GVAKTGLSGAAGSNQ--FFTFVPAK-----TNVFTMSGGTGDAD 550

QY 370 LVITAPNGTXYVNDPFAFPDNNMGR-----NNVENVFINSPOSGTYTIEVQAYNVPV 424

Db 551 LVYK-----LG--SQPTSSYDCKRPEYGENAEVCSFAPQAGTYHWINGKAYSG 599

RESULT 13

A72647

probable surface layer-associated STABLE proteinase APE0607 - Aeropyrum pernix (strain

C:Species: Aeropyrum pernix

C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C/Accession: A72647

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Taka

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A:Reference number: A72450; MUID:99310393; PMID:10382566

A:Accession: A72647

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1331 <XAM>

A:Cross-references: DDBJ:AP000060; NID:95104188; PIDN:BAA79577.1; PID:d1043363; PID:g51

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0607

Query Match 11.6%; Score 260; DB 2; Length 1331;

Best Local Similarity 24.0%; Pred. No. 9.1e-09;

Matches 113; Conservative 67; Mismatches 153; Indels 138; Gaps 19;

QY 8 VKADVAQSSVGLYGGQIVAVADTGLDGRNDSMHE-ARFKITLXYA----- 55

Db 346 VYADLS-TAYLTF---LKALSDTGMTSGEPDPSLDDLSPADETPASYSGEVLARDFG 400

QY 56 -----LGRTNNAN-----DTNGH 68

Db 401 GVNDFSGALAGWTYDWVGLTGESVNLGRLGFDYAGLVLPGLDPQGRWVSLIYDTLAI 460

QY 69 GTHVAGSLTNG-----ANKMAPOANLVPSIIMDSGGGLGGLPSNIQTLE--SQ 117

Db 461 GTSVAITYASRGVNEVNLGYETSLRGVAPGAKI-----AAGG-----SFLINFAAQ 508

QY 118 AFSAG-----ARHTNSWG--AAVNGAYT--TDSRVNDVY-VKNDM 154

Db 509 LFLSGFQDPSPLMWVYTGHEQVQVINSWNGSVIALRGTLGADVDYATIEDIYVASGT 568

QY 155 TILPAAGNEGNGTISAPGTAKNATVGTATE--NLAPSGVSADNINVAQSSRQPTK 212

Db 569 VIVAMKNGGQGYTATTPAGSLITSVGASTLFEDYVFPFGYVPSPGGDVYISMDRSPQ 628

QY 213 DGRITPDVMAAGTILSARSLAPDSFWANHDSKYAVMGSTMAPTVAGNVAQLREHF 272

Db 629 IGVAKPDVNNIGSPAMNG-----VPVLTLGLNGSLAPDIPGTSBATPMSGSVALVISAY 684

QY 273 VKRNGITPKPSLLKALIAAGADVGLGYPNGNQMGKRVTLDKSLNV-----A 319

Db 685 QOAFGKPSGLVYKALIKSTARDTGA--DAFTQSGSQVDVYRAVAVLEGVPIALSTSV 742
QY 320 YVNESALSTSQKATYFTATAGKPLKISLWSDA--PASTTASVTVNDL 368
Db 743 YENVYLLS-----GYSYEFLLAPNPFEDDQITPGVLKPGETAVETVLVTKTL 788

RESULT 14
A35742
aqualysin (EC 3.4.21.-) I precursor - Thermus aquaticus
C:Species: Thermus aquaticus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A35742; S00620; S00324
R:Terada, I.; Kwon, S.T.; Miyata, Y.; Matsuzawa, H.; Ohta, T.
J. Biol. Chem. 265, 6576-6581, 1990
A:Title: Unique precursor structure of an extracellular protease, aqualysin I, with NH-2
A:Reference number: A35742; PMID:9021674; PMID:2182621
A:Accession: A35742
A:Molecule type: DNA
A:Residues: 1-513 <TER>
A:Cross-references: GB:D90108; GB:D90108; GB:J05414; NID:G217171; PIDN:BA41435.1; PID:G
A:Note: The authors translated the codon CTG for residue 470 as Val, and GGT for residue
R:Kwon, S.T.; Terada, I.; Matsuzawa, H.; Ohta, T.
Eur. J. Biochem. 173, 491-497, 1988
A:Title: Nucleotide sequence of the gene for aqualysin I (a thermophilic alkaline serine
A:Reference number: S00620; PMID:8825062; PMID:3286255
A:Accession: S00620
A:Molecule type: DNA
A:Residues: 75-442 <KWO>
A:Cross-references: EMBL:X07734; NID:G48069; PIDN:CA0559.1; PID:G602091
A:Note: Part of this sequence, including the amino and carboxyl ends of the mature prote
R:Matsuzawa, H.; Tokugawa, K.; Hamachi, M.; Mizoguchi, M.; Taguchi, H.; Terada, I.; Kwo
Eur. J. Biochem. 171, 441-447, 1988
A:Title: Purification and characterization of aqualysin I (a thermophilic alkaline serin
A:Reference number: S00324; PMID:88151957; PMID:3162211
A:Accession: S00324
A:Molecule type: protein
A:Residues: 128-170 <MATS>
C:Superfamily: subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine protease
F:1-14/Domain: signal sequence #status predicted <PRO>
F:15-127/Domain: propeptide #status predicted <PRO>
F:128-408/Product: aqualysin I #status experimental <MAT>
F:157-364/Domain: subtilisin homology <SBT>
F:255-257/Region: S1 specificity crevice #status predicted
F:409-513/Domain: carboxyl-terminal propeptide #status predicted <CPR>
F:166,197,349/Active site: Asp, His, Ser #status predicted

Query Match 11.6%; Score 259.5; DB 1; Length 513;
Best Local Similarity 26.4%; Pred. No. 2.8e-09;
Matches 119; Conservative 45; Mismatches 141; Indels 145; Gaps 24;

QY 16 SYGLYQGOIVAVADTGLDTGRNDSMEAPFGKITALY-ALGRNANNDTGHGTHVAG 74
Db 152 TYRATRGVNVYIDIGIRT-----THREFGRRARVGDALG--GNGDCNGHGHVAG 203
QY 75 SVLGNGATNKGAPOANLVFQSIIMD--SSGGLGLGPSNLQTLFSGAFSAGARIHTN---- 128
Db 204 TI---GGVTVGAKANLVAVVLDGNGSGTSGVAVGDVWV-----TNRHRRPAAVA 252
QY 129 ---SWGAAVNGAYTDSRNVDDYRKX---DMTILFAGNEGNGGTISAPGAKNAITVG 183
Db 253 NNSLGGGVSTA-----LDNAVKSIAAGCVYAVVAGNDNNAKCVS--PAFVAEALTVG 304
QY 184 AT--ENLRPSFGSYADININVAQFSRGPDKGRIKPDVMAPTIYLSARSLAPDSFW 241
Db 305 ATTSAPARASFSNYSQCV-----DLFAPGASIPSA-----W 335
QY 242 ANHDSKAYAVMGSTMATPIVAGNVAAQLREHFVKNRGITP---KPSILKALLAGADVGL 298
Db 336 YTSDIATQTINGTSMATPHVAG--VAAL--YLEONPEATPASVASALNGATTGRLSIGS 392

QY 299 GYPNGQMGWRYTLDKSLNVAAYNESALSTSQKATYFTATAGKPLKISLWSDAPAST 358
Db 393 GSPN-----RLLY-----SLSSGSGST-----APCTS 415
QY 359 ---TASVTLVNDLVITAPNGTRY-----VGNDFSAFPNNMDGN--- 397
Db 416 CSYITGSLSGPDYNF---OPNGIYYSPPAGTHRAMLRGPAGTDFDL-YMRWDGSRMLT 471
QY 398 -----NVENVFINSQSGTYTIEVQAVN 420
Db 472 VGSSTGPTSESLSYSGTAGYLMRIYAVS 501

RESULT 15
D75393
serine proteinase, subtilase family - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: D75393
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathavan, J.O.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; M
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; PMID:20036896; PMID:10567266
A:Accession: D75393
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-627 <WHI>
A:Cross-references: GB:AE001990; GB:AE000513; NID:G6459214; PIDN:AAFI1026.1; PID:G64592
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1459
A:Map position: 1

Query Match 11.5%; Score 259; DB 2; Length 627;
Best Local Similarity 25.7%; Pred. No. 3.9e-09;
Matches 105; Conservative 54; Mismatches 178; Indels 72; Gaps 15;

QY 18 GIVYQGOIVAVADTGLDTGRNDSMEAPFGKITALYAGR-----TNNANDTNGHGHV 72
Db 161 GFGQNRIRIGHDSGIDPS-----HPELAGKVAALQZFNGBDRVSGPHDTTHGHIT 214
QY 73 AGSVLNGATNKGAPOANLVFQSIIMDSSG-----LGLGPSNLQTLFSGAFSAGARIHT 127
Db 215 AGLLVY---SKVGAAPGAKVISALVFPNNGTPAQVIAGWQVLDPDNNADTDGDVNVN 271
QY 128 NSNGAANGAYTDSRNVDDYRKNDMTILFAGNEGPNCGTISAPGAKNAITVATEN 187
Db 272 MSIG--IPGTWNEPIVFNMM--KAGVVPVAFAGNFPAGSGTSGPNTLPQALGVGA--- 325
QY 188 LRPSFGSYADININVAQFSRGPDK--DGR-----KPDVMAPTIYLSARSLAPDSFW 241
Db 326 -----VDSNGQVASFSSRGVAVAMOGELSGVFTFPDIAAPGVNITS----- 365
QY 242 ANHDSKAYAVMGSTMATPIVAGNVAAQLREHFVKNRGITP---SLKPAALAGADVG 297
Db 366 TVRNGVQANSGSSQASPIITAGAAVL-----LSAKPGASVDAIKNAALFTSASNAS 416
QY 298 LGVYNGQMGWRYTLDKSLNVAAYNESALSTSQKATYFTATAGKPLKISLWSDAPAS 357
Db 417 A--KNNNVGFGQISIPALQKGV---STGTBPQPPRPAPPTPTPTPTTPAPAP 471
QY 358 TIASVTLVNDLVITAPNGTRYV--NDPSAFPNNMDGNVENVF 403
Db 472 PAPAAPETGPAGYTLCAVEGRCSGAROKOVAGTGRWISGTSTDDTF 520

Search completed: March 31, 2004, 16:10:16
Job time : 13.3506 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 7.51154 Seconds
(without alignments)

3008.498 Million cell updates/sec

Title: US-09-985-689a-2

Sequence: 1 NDVAGIVADVAQSSYGLY.....EVQAVNVPGPNESLAVN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	526.5	23.5	1743	1 TAGC_DICDI	Q23868 dictyostell
2	490.5	21.8	1905	1 TAGS_DICDI	P54683 dictyostell
3	304.5	13.6	1398	1 PLS_PIRFU	P72186 pyrococcus
4	302.5	13.5	806	1 SUBV_BACSU	P29141 bacillus su
5	278	12.4	420	1 SUBV_BACSU	P28842 bacillus su
6	277	12.3	580	1 EXPR_XANCP	P23114 xanthomonas
7	259.5	11.6	513	1 AQLI_THBAO	P08594 thermus agu
8	250.5	11.2	534	1 PROA_VIBAL	P16588 vibrio algi
9	250.5	11.2	402	1 ALP_CEPAC	P29118 cephalospor
10	249.5	11.1	894	1 WPR4_BACSU	P54423 cephalospor
11	246	11.0	269	1 SUBS_BACLE	P29600 bacillus le
12	246	11.0	380	1 ELVA_BACAO	P27693 bacillus al
13	246	11.0	380	1 ELVA_BACAO	P41362 bacillus al
14	245.5	10.9	401	1 THES_BACSP	Q45670 bacillus sp
15	245	10.9	269	1 PRTW_BACSP	Q99405 bacillus sp
16	240	10.7	378	1 ELVA_BACSP	P20724 bacillus sp
17	235	10.5	269	1 SUBS_BACLE	P29599 bacillus le
18	234.5	10.4	379	1 SUBT_BACLI	P00780 bacillus li
19	234.5	10.4	382	1 SUBT_BACAM	P00782 bacillus am
20	232.5	10.4	1052	1 MS1P_HUMAN	Q14703 homo sapien
21	232.5	10.4	1181	1 SCAL_STRPY	P58099 streptococc
22	232	10.3	331	1 ISP_BACCS	Q92440 bacillus cl
23	231.5	10.3	1052	1 MS1P_CRICR	Q92448 cricetulus
24	231.5	10.3	1052	1 MS1P_MOUSE	Q9wcz2 mus musculu
25	231.5	10.3	1052	1 MS1P_RAT	Q9wcz3 rattus norv
26	229.5	10.2	381	1 SCAL_STRPY	P15926 streptococc
27	229	10.2	1167	1 SUBT_BACNA	P04186 bacillus su
28	227.5	10.1	381	1 SUBT_BACSU	P00718 bacillus su
29	225.5	10.0	275	1 SUBT_BACPU	P00718 bacillus su
30	225.5	10.0	381	1 SUBT_BACSA	P29142 bacillus su
31	225.5	10.0	381	1 SUBT_BACST	P29142 bacillus su
32	224	10.0	404	1 SMP1_MAGPO	Q9y778 magnetospor
33	219.5	9.8	274	1 SUBD_BACLI	P00781 bacillus li

34	219.5	9.8	645	1 SUBB_BACSU	P16396 bacillus su
35	218	9.7	279	1 THEV_THREU	P04072 thermoactin
36	217.5	9.7	361	1 ELVA_BACAD	P41363 bacillus ha
37	217	9.7	293	1 PRIT_TRIAL	P20015 trichachiu
38	217	9.7	422	1 TKSU_PYRKO	P58502 pyrococcus
39	214	9.5	326	1 ISP_FABPO	P29139 paenibacill
40	213	9.5	409	1 ALP_TRIHA	Q03420 trichoderma
41	211	9.4	319	1 ISPI_BACSU	P11018 bacillus su
42	211	9.4	467	1 ISP6_SCHPO	P40903 schizosacch
43	210.5	9.4	388	1 CUDP_MERAN	P29138 metathizium
44	209.5	9.3	530	1 HLY_HALI7	P29143 halophilic
45	208.5	9.3	1902	1 P2P_LACCA	Q02470 lactobacill

ALIGNMENTS

```

RESULT 1
TAGC_DICDI          STANDARD;          PRT; 1743 AA.
AC  Q23868;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  Prestalk-specific protein tagc precursor (EC 3.4.21.-).
CN  TAGC.
OS  Dictyostelium discoideum (Slime mold).
OC  Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX  NCBI_TaxID=44689;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=AX4;
RX  MEDLINE=97140317; PubMed=8986798;
RA  Shuisky G., Escalante R., Loomis W.F.;
RT  "Developmental signal transduction pathways uncovered by genetic
RT  suppressors.";
RL  Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).
CC  -!- FUNCTION: Interleukin communication via tagc may mediate
CC  integration of cellular differentiation with morphogenesis (by
CC  similarity).
CC  -!- SIMILARITY: In the N-terminal section; belongs to peptidase family
CC  S8.
CC  -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
CC  TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC  -!- SIMILARITY: STRONG, TO TAGB.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; U60086; AAB0331.1; -.
DR  PIR; T18279; T18279.
DR  DictyBase; DDB0001795; tagc.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001140; ABC_TM_transp.
DR  InterPro; IPR003439; ABC_transporter.
DR  InterPro; IPR000209; Peptidase_S8.
DR  Pfam; PR00664; ABC_membrane; 1.
DR  Pfam; PR00005; ABC_tran; 1.
DR  Pfam; PF00082; Peptidase_S8; 1.
DR  PRINTS; PR00723; SUBTILISIN.
DR  SMART; SMO0382; AAA; 1.
DR  PROSITE; PS50929; ABC_TMLF; 1.
DR  PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR  PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR  PROSITE; PS00135; SUBTILASE_ASP; FALSE_NEG.
DR  PROSITE; PS00137; SUBTILASE_HIS; 1.
DR  PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
DR  HydroBase; Serine protease; ATP-binding; Transport; Transmembrane;
KW  P00781 bacillus li

```

```

KW Signal.
FT CHAIN 1 27
FT CHAIN 28 1743
FT CHAIN 316 642
FT CHAIN 1450 1687
FT TRANSMEM 962 982
FT TRANSMEM 1027 1047
FT TRANSMEM 1072 1092
FT TRANSMEM 1157 1177
FT TRANSMEM 1260 1280
FT TRANSMEM 1288 1308
FT ACT_SITE 325 325
FT ACT_SITE 372 372
FT ACT_SITE 637 637
FT NP_BIND 1485 1492
FT NP_BIND 42 46
FT DOMAIN 94 103
FT DOMAIN 643 646
FT DOMAIN 733 741
FT DOMAIN 786 792
FT DOMAIN 1337 1340
FT DOMAIN 1346 1352
FT DOMAIN 1353 1357
FT DOMAIN 1358 1364
FT DOMAIN 1381 1386
FT DOMAIN 1707 1729
FT CARBOHYD 390 390
FT CARBOHYD 536 536
FT CARBOHYD 547 547
FT CARBOHYD 614 614
FT CARBOHYD 689 689
FT CARBOHYD 735 735
FT CARBOHYD 741 741
FT CARBOHYD 776 776
FT CARBOHYD 832 832
FT CARBOHYD 887 887
FT CARBOHYD 1251 1251
FT CARBOHYD 1385 1385
FT CARBOHYD 1454 1454
FT CARBOHYD 1704 1704
SQ SEQUENCE 1743 AA; 194145 MW; 12DB363E2F729839 CRC64;

Query Match 23.5%; Score 526.5; DB 1; Length 1743;
Best Local Similarity 27.8%; Pred. No. 9,6e-27;
Matches 164; Conservative 82; Mismatches 152; Indels 191; Gaps 22;

QY 19 LYGGGQIVAVADTGLDTGR--NDS-----SMHEAFRGKITALYALGRTNANNNDNGH 68
DB 314 LRGGQQLIADIADTGLDQSHCFPSDSKYPILPNSVNLNHR-KVVTYITTTSDSDSKVDGH 372
QY 69 GTHVAGSVLG-----NGATNKGMAPQANLVPFSIMDSGGVGLPSPNQLTFSQAFA 121
DB 373 GTHICGAAGTPEDSSVNISSFGLARDAKIAFPDLSSGSSSLTP-PSDLKQLYQPLVXD 431
QY 122 GARHHTNSWGA---AVNGAYTDSRVVDYVRKN-DMTILFAAGNEGPGGITS--APG 174
DB 432 GARHCHDSWGSVSEGYTGSYSTASIDFLFTHPFIILRAAGN---NEQYLSLITGS 486
QY 175 TAKNAITVGAATENR-----PSFGSYADNI----- 199
DB 489 TAKAVITVGAHQTHENYLTIDGPNYINYGSSVDINQELICDFPSRCYNTTAQCCLESNA 548
QY 200 -----NRVAFSSRGPTKXGRINQEDVMAFGTYILSA 230
DB 549 TTGLASCCPTLLRKSVIDAANTQPLLYNENNICFSSSKGPTHGRMKPALVAGREVITSA 608
QY 231 RSSLA-----PDSFPAHNDSKAYMGTSMAPIYAGVAAQLEH-----FYK 274
DB 609 RSNQANNTTDOCGDGL-PTNTAALIA-LSGTSKATSPFAAAATTLIRQYLVGDYPTGSIVE 666
QY 275 NRGITPKPSLLKALILAGA-----ADVGLGYFNGN-----QCMGRTVLD 313

```

```

DB 667 SNKQPTGSLKALMINNAQLNGTFQLTSSITPSNOVFENFAGASLVQMGAIRMS 726
QY 314 KSLNVAAYNESS-----ALSTQKATYTF----- 338
DB 727 NMLHVNANNNNNNNNKTSIDGITKFDGIGDLRLVKRNQWKEBLSIGQNTSYCYFYKPS 786
QY 339 ---ATAGKELK---ISLVSDAPASTASVTLVNDLVLVITAPRGTRYVGNDFAPDNN 392
DB 787 SSSNSGNNIRRVAVATLWIDPPEYAGKKNLVNNDLVT-----YRDNSTIYSN 840
QY 393 WDGR-----NVVENVPINSQSGTYTIEVQAYVVPQPNFS 429
DB 841 QGSSFLGLAFTQDTLNNVEGIVANPTEPMYRPMVAGTVMPGQPNFS 889

RESULT 2
TAGB_DICDI STANDARD; PRT; 1905 AA.
ID TAGB_DICDI
AC P54683;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prestalk-specific protein tagb precursor (BC 3.4.21.-).
GN TAGB.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4.
RX MEDLINE=95562903; PubMed=7744252;
RA Shaulsky G., Kuspa A., Loomis W.F.;
RT "A multidrug resistance transporter/serine protease gene is required for prestalk specialization in Dictyostelium."
RL Genes Dev. 9:1111-1121(1995).
CC -1- FUNCTION: Intercellular communication via tagb may mediate integration of cellular differentiation with morphogenesis.
CC -1- SIMILARITY: In the N-terminal section; belongs to peptidase family 88.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -1- SIMILARITY: STRONG, TO TAGC.
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk.)
CC
CC EMBL, U20432; AAA62212.1; -.
CC PIR, T18267; T18267.
CC
CC MEROPS: S08.CPW; -.
CC DictyBase; DB0001964; tagb.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR001140; ABC TM transp.
CC InterPro; IPR003439; ABC transporter.
CC InterPro; IPR002099; Peptidase_S8.
CC Pfam; PF00664; ABC_Membrane; 1.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC ProDom; PD000006; ABC transporter; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS50929; ABC_TM1F; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane; Signal.
KW

```

```

FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 1905 PRESTALK-SPECIFIC PROTEIN TAG.
FT DOMAIN 376 700 PROTEASE.
FT DOMAIN 1518 1756 ABC TRANSPORTER.
FT TRANSMEM 1011 1031 POTENTIAL.
FT TRANSMEM 1076 1096 POTENTIAL.
FT TRANSMEM 1121 1141 POTENTIAL.
FT TRANSMEM 1210 1230 POTENTIAL.
FT TRANSMEM 1309 1329 POTENTIAL.
FT TRANSMEM 1332 1352 POTENTIAL.
FT ACT SITE 387 432 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 432 432 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 695 695 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1553 1560 ATP (POTENTIAL).
FT DOMAIN 63 67 POLY-GLN.
FT DOMAIN 95 104 POLY-ASN.
FT DOMAIN 107 134 POLY-ASN.
FT DOMAIN 311 321 POLY-SER.
FT DOMAIN 833 837 POLY-SER.
FT DOMAIN 838 844 POLY-GLY.
FT DOMAIN 871 876 POLY-LEU.
FT DOMAIN 1012 1015 POLY-ILE.
FT DOMAIN 1386 1389 POLY-GLY.
FT DOMAIN 1398 1404 POLY-GLY.
FT DOMAIN 1445 1450 POLY-ASN.
FT DOMAIN 1765 1779 POLY-ASN.
FT DOMAIN 1782 1785 POLY-SER.
FT DOMAIN 1807 1812 POLY-PRO.
FT DOMAIN 1813 1860 POLY-PRO.
FT DOMAIN 1872 1878 POLY-PRO.
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1172 1172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1522 1522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1658 1658 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1905 AA; 212518 MW; B8E223FA8B9AE13C CRC64;

```

Query Match 21.8%; Score 490; DB 1; Length 1905;

Best Local Similarity 28.0%; Pred. No. 2,7e-24; Indels 184; Gaps 23;

Matches 162; Conservative 74; Mismatches 159; Indels 184; Gaps 23;

```

QY 19 LVGQGOIVAVADTGLDPTGR---NDS-----SMHEAFGKITALYALGRTNNANDTNGH 68
DB 376 LRGKGLISTADTGLDGSHPFSDSKYPIPFNVNENHRYVT---YIYHNDYVANGH 432
QY 69 GTHVAGSVLANG-----ATNKGAPOANLVFQSIIMDSGGIGGL-PSNLQTLFSQAFS 120
DB 433 GTHVCGSAAGTPEDSSWAISSFGLATDAKIAFYDL--SSGSSEPTPPEDYSQMYKFLVD 490
QY 121 AGARHITSWGA---AVNGAYTTDSRNVDDYRK-NDMTILFFAAGNEGNGGTSAPGT 175
DB 491 AGARHVDSDSGSVSLGQYGYGVDDAGIDAFIYEPERSILRAAGN-NELTASILAQAT 549
QY 176 AKRAITVATENTLRPSFGS-----YADNT----- 199
DB 550 AKRAITVGAEOIYAVVYVSDALEYVDFSDANFORPCLFDKRYCNYTTAKCSESVNYGK 609
QY 200 -----NHVAQFSSRGPTDGRIRKIPVMAFGYIISARS 233
DB 610 LQLOCPASIKONASDSTTPOFYENNNMGSFSSKPTTHDGRUKPIYVAPGEYIISARN 669
QY 234 -----LAPDSSWANHSKRYAMGTSNATPIVAGNVACLREHF-----YKONG 277
DB 670 GENSTQCGDGSGL--PIANGIMISGTSWATPLATAATILIQYLVLDGFGFPGSEVEENK 727
QY 278 ITRKPSLTKAALJAGADVGLG-----PENGQSGRATLDSKLV 318
DB 728 LITPTGISLKLMTNNALNGTYFWBASSTNPENALFEQINGANLILQNGALNKN--NW 784
QY 319 AYNNESS-----ALSTSOKAT-----YTF-----ATAG 342

```

```

DB 765 LYVASSNPTSPRWIGIGIGKQKATEWKEDLSGLKNSKYCFYKPSSSSSSGSGGGG 844
QY 343 KP-LKILVWSDAPASATTAIVTLVNDL-----VITAN--GTRVGNDFAPFD 350
DB 845 TPRVATLVWTDPSYSGAFNLVNNLDDLNSDDSLITIGSGSLQPAKVAP-- 902
QY 391 NMDGRNVENVPINSPOSGTYTIEVOYVNVPGVQNF 429
DB 903 ---DTLNVGIIINPTKANVYKFTIAGTVNPIGPQF 938

RESULT 3
PIS PYRPU STANDARD; PRT; 1398 AA.
AC P72186;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pyrolysin precursor (EC 3.4.21.-).
GN PLS OR P70287.
OS Pyrococcus furiosus.
OC Pyrococcus.
CC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 150-184, AND CHARACTERIZATION.
RC STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=96355370; PubMed=8702780;
RA Voorhorst W.G.B., Eggen R.I.L., Geerling A.C.M., Platteeuw C.,
RA Sieszen R.J., de Vos W.M.;
RT "Isolation and characterization of the hyperthermostable serine
RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
RT Pyrococcus furiosus.";
RL J. Biol. Chem. 271:20426-20431(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weis R.B., Dunn D.M., Robb P.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION, AND 3D-STRUCTURE MODELING.
RX MEDLINE=21079021; PubMed=11210516;
RA de Vos W.M., Voorhorst W.G.B., Dijkgraaf M., Kluskens L.D.,
RA Van der Oost J., Sieszen R.J.;
RT "Purification, characterization, and molecular modeling of pyrolysin
RT and other extracellular thermostable serine proteases from
RT hyperthermophilic microorganisms.";
RL Mech. Enzymol. 330:383-393(2001).
CC -I- FUNCTION: Has endopeptidase activity toward caseins, casein
CC fragments including alpha-S1-casein and synthetic peptides.
CC -I- SUBCELLULAR LOCATION: Cell envelope associated.
CC -I- PIM: LHM pyrolysin seems to be produced by autoprolytic
CC activation of HM pyrolysin.
CC -I- PIV: Glycosylated.
CC -I- MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.
CC -I- SIMILARITY: Belongs to peptidase family S8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: U55835; AAB09761.1; -
DR EMBL: AE010153; AAL60411.1; -
DR PIR: T28159; T28159.
DR HSRF: Q45670; ID51.
DR MEROPS: S08.100; -.
DR InterPro: IPR000209; Peptidase_s8.

```

DR InterPro: IPR007280; PPC.
 DR Pfam: PF00082; peptidase_S8; 1.
 DR Pfam: PF04151; PPC; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE ASP; 1.
 DR PROSITE: PS00137; SUBTILASE HIS; 1.
 DR PROSITE: PS00138; SUBTILASE SER; 1.
 KW Hydrolyase, Serine protease; Glycoprotein; Zymogen; Signal;
 KM Complete proteome.
 FT SIGNAL 1
 FT PROPEP 26
 FT CHAIN 27
 FT ACT_SITE 150
 FT ACT_SITE 179
 FT ACT_SITE 365
 FT ACT_SITE 590
 FT CARBOHYD 152
 FT CARBOHYD 152
 FT CARBOHYD 222
 FT CARBOHYD 228
 FT CARBOHYD 240
 FT CARBOHYD 240
 FT CARBOHYD 257
 FT CARBOHYD 262
 FT CARBOHYD 298
 FT CARBOHYD 327
 FT CARBOHYD 406
 FT CARBOHYD 651
 FT CARBOHYD 651
 FT CARBOHYD 663
 FT CARBOHYD 739
 FT CARBOHYD 792
 FT CARBOHYD 893
 FT CARBOHYD 908
 FT CARBOHYD 917
 FT CARBOHYD 929
 FT CARBOHYD 1048
 FT CARBOHYD 1056
 FT CARBOHYD 1084
 FT CARBOHYD 1117
 FT CARBOHYD 1133
 FT CARBOHYD 1140
 FT CARBOHYD 1148
 FT CARBOHYD 1208
 FT CARBOHYD 1233
 FT CARBOHYD 1237
 FT CARBOHYD 1332
 FT CONFLICT 607
 FT CONFLICT 881
 FT CONFLICT 881
 SQ SEQUENCE 1398 AA; 154474 MW; 355D873A27D56582 CRC64;
 Query Match 13.68; Score 304.5; DB 1; Length 1396;
 Best Local Similarity 26.38; Pred. No. 2,7e-12;
 Matches 138; Conservative 57; Mismatches 177; Indels 153; Gaps 22;
 21 GGGGIVAVADTGLDPTGRNDS-----SMHEAFKGTALYALGRNNAN----- 63
 301 GNGGDIYVYTDYDFTDEVPFGQYVTVDAVFSSYGGPLNYVLAIEDPGEYVAFGM 360
 64 DTNGHGHNAGVSLGNGATN-----KGNAPQAN 91
 361 DGHGHGHNAGVSLGNGATN-----KGNAPQAN 91
 92 LVPQSLNDSSGGTGLGPNLQTFSSQAFSAGARIHNSWGAANVAYT--TDSRN--VDD 147
 421 IMAIRVY-RSDGRGSMWDIEGM-TYATAGADIVSISLGG--NAYLGDTPESVAVDE 476
 148 YVAKNDITLFAAGNENGGTISAPGTAKATVGTENTLRPSFSYVD----- 197
 477 LTKYGVVFIAGNBPAGINIVGSPVATKATLVGAA--VPINAGVYSQALGFPDYVG 535
 198 -----NINHAQFSSRGPTDGRIRKPDVMAPGYIISARESLAPDSSFWANHDSKYA 249
 536 FYFPAAYTNV-RIAFSSRGPRIRIGEKPNVVAAGYIYSSLPMMIGADF----- 585
 250 YMGTSWATPIVAGNVAQLREHFVKNCGITPKPSILKALNAGAA-----DVGGLGYPNG 303

DB 586 -MSGTSWATPIVAGNVAQLREHFVKNCGITPKPSILKALNAGAA-----DVGGLGYPNG 303
 QY 304 NOGNGRVTLDKSLNVAAYNESSALSTSOKATYFTATNGPKLKSILVWSDPASTTA--- 360
 DB 644 DGHGHGHNAGVSLGNGATN-----KGNAPQAN 91
 QY 361 SVTLVNDLDVITFPN-----GTRYVNDPFSAPD-----NNW-----DG-----RNVE--- 400
 DB 686 GVDVIRGLVYARNISIPDIVEMHIKIVGDTERTREIYATEFEMIKPFSVGLNNTFEVL 745
 QY 401 -----NVFINSPOSGTYTTEVOAYNPVGPONFS 429
 DB 746 RVXVDVEGLEGLVGRILIDDP--TPVEDELINIVIPERFT 788
 RESULT 4
 SUBV_BACSU STANDARD; PRT; 806 AA.
 AC P29141;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Minor extracellular protease vpr precursor (EC 3.4.21.-).
 GN VPR OR IPA-45R OR BSU38090.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 161-195.
 RX MEDLINE=92041574; PubMed=193892;
 RA Sliema A., Ruto G.A. Jr., Theriault K.A., Dwyer M., Wilson S.W.,
 RA Pero J.;
 RT "Cloning and characterization of the gene for an additional
 RT extracellular serine protease of Bacillus subtilis";
 RT J. Bacteriol. 173:6869-6885(1991).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=168;
 RC MEDLINE=95020537; PubMed=7934828;
 RX Glaeser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
 RA Hullu M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
 RA Prescan E., Santana M., Schneider B., Schweizer J., Vertes A.,
 RA Rapoport G., Danchin A.;
 RT "Bacillus subtilis genome project: cloning and sequencing of the 97
 RT kb region from 325 degrees to 333 degrees.";
 RT Mol. Microbiol. 10:371-384(1993).
 [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=168;
 RC MEDLINE=96044033; PubMed=9384377;
 RX Kunst F., Ogatawa N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berteiro M.G., Bessieres P., Bolotin A., Borhett S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.D., Daniel R.A.,
 RA Dehioz F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Brittan K.D., Errington J., Fabre C., Ferrazi E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi E.J., Grandi G.,
 RA Ghu S.Y., Glaeser P., Goffeau A., Gohgichly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Halesch J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holappel S., Hosono S., Hullu M.F., Itaya M., Jones J.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapidas A., Lardinois S., Lauber J., Lazarevic C.,
 RA Lee S.M., Levine A., Liu H., Maeda S., Manel C., Medigue C.,
 RA Medina N., Melisado R.P., Mizuno M., Moser D., Nakai S., Nodack M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
 RA Pao D., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
 RA Priesen E., Puig C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Sciffone F.,
 RA Sekiguchi J., Sekowska A., Serrif S.J., Serrif P., Shih B.S., Solido B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru T.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosaio V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambolt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yanane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Darchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*,"
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: NOT REQUIRED FOR GROWTH OR SPOULATION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PTM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M76590; AAA22881.1; -
 DR EMBL: X73124; CA451601.1; -
 DR EMBL: Z39123; CAB15835.1; -
 DR PIR: A41341; A41341.
 DR HSSP: P00783; 2S8T.
 DR MEROPS: S08.00A; -
 DR Subtilisin, BGI0591; VPR.
 DR InterPro: IPR003137; PA.
 DR InterPro: IPR000209; Peptidase_S8.
 DR InterPro: IPR009020; Protease_inhib.
 DR Pfam: PF02225; PA; 1.
 DR Pfam: PF02225; PA; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILISIN_ASP; 1.
 DR PROSITE: PS00137; SUBTILISIN_HIS; 1.
 DR PROSITE: PS00138; SUBTILISIN_SER; 1.
 KM Hydrolyase; Serine protease; Zymogen; Signal; Complete proteome.
 FT SIGNAL 1 28
 FT PROPEP 29 160
 FT CHAIN 161 806
 FT ACT_SITE 189 189
 FT ACT_SITE 233 233
 FT ACT_SITE 534 534
 SQ SEQUENCE 806 AA; 85608 MW; P984E3BF0869DDC CRC64;
 Query Match 13.5%; Score 302.5; DB 1; Length 806;
 Best Local Similarity 23.2%; Pred. No. 1,8e-12;
 Matches 127; Conservative 58; Mismatches 161; Indels 201; Gaps 17;
 QY 18 GLYGQGIYVAADTGLDTR-----NDSSVHEAFPRKITALVALGRTN 60
 DB 177 GYTGKGIKVALIDTGEVNHPRDKKFGYGVDFVNDYDKRPTG-----D 225
 QY 61 NAADTNGHGHVAGSLVNGATKKAAPQANLVFQSIIMSSGLGULPSNLTLSQAPS 120
 DB 226 PRGEATDGHVAGVLAANG-TIKGVAPEATLLAVRLPGG--SGTENVAVGERAVQ 282
 QY 121 AGARHTNSWGAANVAYTTDSRNVDYRKNDMTILFAAGNEGPNGTISAPTAQAI 180
 DB 283 DGDVWNLSIGNSLNNPDWATSTAL-DWAMSEGVAVYVNSNGSNGMVGSPGISRAI 341
 QY 181 TVGATE---NLPSFGSY----- 195
 DB 342 SVGATOLPLNEVAVTGESYSASAKVGNKEDDVKALNNKEVELVEAGIGEAQDFGKDL 401
 QY 196 -----ADNI----- 199
 DB 402 GKVAVVKRGSIAVVDKADNKKAGATGMVYNNLSGEIAPNPGMSVPTIKSLDEGKL 461
 QY 200 -----NHVAFSSRGPTKD-GIKFDVAPAGTYTILSASSSLAP 236
 DB 462 VSALKAGETKTKTLTVSKALGEQVADFPSSRGVMDTWIKPDISAPGVNIVSTIPTHDP 521

QY 237 DSSFWANHDSKAYVMGSTWATPIVAGNVAQLEHFWKRGITPKPSL--IKALTAGAA 294
 DB 522 D-----HPYTGSKQSTSNASPHILKAIVAVIKQ-----AKPKASVEGIRKALINTAY 568
 QY 295 DV-----GLGYPNGNQGNGVYTLDKSLNVAVYNESSALSTQKATYFTATAGKPLKISLV 350
 DB 569 TLKSDSEVYPHNAQAGASARI---MVA--IKADSLVSPGSYSGFTLKENGNETKNETF 623
 QY 351 WSDPASTTASVTLVNDLVIAPNCTRVGNVDFGAPFNMDGNNVENVFINPSQG 410
 DB 624 TIENQSSIRKSYTL-----EYS--FVSGSISTSGISRVVIPAHTG 662
 QY 411 TYTIEVQ 417
 DB 663 KATKVK 669
 RESULT 5
 SUBT_BAC93 STANDARD; PRT; 420 AA.
 AC P28842;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Subtilisin precursor (EC 3.4.21.62).
 GN SUB1.
 OS *Bacillus* sp. (strain TA39).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=29336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92256481; PubMed=1581352;
 RA Narinx E., Davail S., Feller G., Gerday C.;
 RT "Nucleotide and derived amino acid sequence of the subtilisin from
 RL the antarctic psychrotroph *Bacillus* TA39,"
 RL Biochim. Biophys. Acta 1131:111-113(1992).
 CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 CC it catalyzes the hydrolysis of proteins and peptide amides;
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 CC for peptide bonds, and a preference for a large uncharged residue
 CC in P1. Hydrolyzes peptide amides.
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (potential).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: Still active at temperatures close to 0 degrees
 CC Celsius; it has a marked heat lability.
 CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 CC sporulation, and many mutations which block sporulation at early
 CC stages affect expression levels of subtilisin. However, subtilisin
 CC is not necessary for normal sporulation.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62369; CA444227.1; -
 DR PIR: S23407; S23407.
 DR HSSP: Q99405; 1WPT.
 DR MEROPS: S08.00A; -
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILISIN_ASP; 1.
 DR PROSITE: PS00137; SUBTILISIN_HIS; 1.
 DR PROSITE: PS00138; SUBTILISIN_SER; 1.
 KM Hydrolyase; Sporulation; Serine protease; Zymogen; Metal-binding;
 KW Calcium-binding; Signal.
 FT SIGNAL 1 26
 FT SIGNAL 26
 POTENTIAL.

```

FT PROPEP 27 111 POTENTIAL.
FT ACT_SITE 112 420 SUBSTITUTIN.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 182 182 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 360 360 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT METAL 115 115 CALCIUM (POTENTIAL).
FT METAL 154 154 CALCIUM (POTENTIAL).
SQ SEQUENCE 420 AA; 44086 MM; AE4F121BD32B266C CRC64;

Query Match 12.4%; Score 278; DB 1; Length 420;
Best Local Similarity 33.2%; Pred. No. 3, 1e-11;
Matches 88; Conservative 38; Mismatches 95; Indels 44; Gaps 13;

QY 21 GGGQIVAVADTGLDTRGNDSSMHEAFRGKITL--VALGPT--NNANDNGCTVAGS 75
Dd 136 GGGININAVDITGVNTN-----HDLNANVEOCKDFVGTYYTNNSCDTRQGGHVAAS 189
QY 76 VLNGATNK--GNAPQANLVFQSIM--DSSGGLGGLPSNLQTLFQASAGARIMTN-S 129
Dd 190 ALADGGTGNVGVVADIMAYKVLGDGSGYADIDIAAIRAGDAQATLNTKVIIMS 249
QY 130 WGAAVNAYTTDSNVVDYVRKNDMTLLFAAGNEPRNGCTISAPGAKNATIGATENR 189
Dd 250 LGSSGSSLLTNNAN--YSYNGVLLTAAAGNSGPGQSIGYPGALVNAVAALLEN-K 305
QY 190 PSFGSYADNINHAQFSSRGPT-KDG-----RIKPDVMAPGTIILSARSLAPDSSFWAN 243
Dd 306 VENGTI-----RVADFSRGSWTDGDIYAIQKGDVEISAGALVST-----W--- 348
QY 244 HDSKYAYMGTSMTPTVAAGNVAQL 268
Dd 349 PDGGYATITSGTSMASPHAGLAKI 373

RESULT 6
EXPR_XANCP STANDARD; PRT; 580 AA.
ID EXPR_XANCP
AC P331.4;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Extracellular protease precursor (EC 3.4.21.-).
GN XCC0851.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadales; Xanthomonas.
CX NCBI_taxid=340;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90251253; Pubmed=2187155;
RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Daniels M.J.;
RT "A multipurpose broad host range cloning vector and its use to
RT characterise an extracellular protease gene of Xanthomonas campestris
RT pathovar campestris";
RL Mol. Gen. Genet. 220:433-440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2022145; Pubmed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canavari F., Cardoso J., Chambeiro F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Mendes J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.C., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,

```

```

RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
CC -|- SUBCELLULAR LOCATION: Secretd.
CC -|- SIMILARITY: Belongs to peptidase family S8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, X51635; CA935962.1; -.
CC EMBL, AE012184; AA040166.1; -.
CC PIR, S11890; S11890.
CC HSSP; P00782; 2SPT.
CC MEROPS; S08 UPA; -.
CC InterPro; IPR000209; Peptidase_S8.
CC InterPro; IPR007280; PFC.
CC InterPro; IPR009020; Protease_inhib.
CC Pfam; PF00082; Peptidase_S8; 1.
CC Pfam; PF04151; PFC; 1.
CC PRINTS; PR00723; SUBSTITIN.
CC PROSITE; PS00136; SUBSTITIN_ASP; 1.
CC PROSITE; PS00137; SUBSTITIN_HIS; 1.
CC PROSITE; PS00138; SUBSTITIN_SER; 1.
CC KMW Hydroxylase; Serine protease; Zymogen; Signal; Complete proteome.
CC FT SIGNAL 1 32
CC FT PROPEP 33 7136
CC FT CHAIN 7137 580
CC FT ACT_SITE 177 177
CC FT ACT_SITE 237 237
CC FT ACT_SITE 409 409
CC FT ACT_SITE 425 425
CC FT DISULFID 315 352
CC FT DISULFID 450 454
CC SQ SEQUENCE 580 AA; 57228 MM; 8C9A2CAB4E7F47CB CRC64;

Query Match 12.3%; Score 277; DB 1; Length 580;
Best Local Similarity 26.6%; Pred. No. 5, 5e-11;
Matches 126; Conservative 53; Mismatches 150; Indels 144; Gaps 22;

QY 21 GGGQIVAVADTGL-----DTGRNDSMHEAFRGKITLVALGRTNNAND----- 64
Dd 168 GSGTVAVAYDITGIRSHDLNANILAGYDPSDATTAIDGNGRNSNADBDGWTAAABCGA 227
QY 65 -----TNGHOTHYAGSVLNGATNKGMAPQ-----NLVQSIMDS 100
Dd 228 GIPAASSWHGTHTAGVAAVNTTGTAGTAAGAKVYPVRLVIGKCGSLSDADLVMA 287
QY 101 SGG-LGGLPSNLQ--TFQAFSAGARIHTNSGAAVNGAYTTDSRVVDYVRKNDMTIL 157
Dd 288 SGGTVSGIPANANAEVYINNSLGGGSCSTTQN-AINGAVSNGT-----TVV 334
QY 158 FAAGNEBNGGTIAPGTAKNATVGTEN--LRPSGSAADNINHAQFSSRGPTDCGR 215
Dd 335 VAAANDASVNG-SLPANCANVIAVAATTSAGAKASVSNGTGI----- 377
QY 216 IKPVMVAPGTIILSARSS--LAPDSSFWANHDSKYAYMGTSMTPTVAAGNVAQLREHFV 273
Dd 378 ---DVSAFGSILSTLNSGTTTPS-----ASYASYNSTSMASPHVAVGVALVQS--V 425
QY 274 KNRGITPK--PSLIK--AALAGADVGLG-----PENGQSGRVTLDK 314
Dd 426 APTALTAAVAVETLLKNTARLPAGSGCGAGIVNDAVTAALINGSGGSGGAGNTLTN 485
QY 315 SLNAAVYNESSALSTQKATYTFATPGKPLKSLVWSDAPASTTASVTL---VNDDLV 371
Dd 486 GTPVT-----GIGATGALNVTITVPAG-----SGTLVTTVSSGSGADADLV 527

```

QY 372 ITAPNGRYVNDSPAFDNNMDGR-----NNVNFINSPOSGTYTTEVOAYN 420
 DB 528 VRAQ-----SAPTDAYTCRPPRSNATCCTITAP-SOTYYVRLKAYS 569

RESULT 7
 AQLI_THEAQ STANDARD; PRT; 513 AA.
 ID AQLI_THEAQ
 AC P08594;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Aqualysin I precursor (EC 3.4.21.-).
 GN PSTI.
 OS Thermus aquaticus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Thermus.
 NCBI_TaxId=271;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 15-23.
 RC STRAIN=YT1;
 RX MEDLINE=9021674; PubMed=2182621;
 RA Terada I., Kwon S.-T., Miyata Y., Matsuzawa H., Ohta T.;
 RT "Unique precursor structure of an extracellular protease, aqualysin
 I with NH2- and COOH-terminal pro-sequences and its processing in
 Escherichia coli.";
 RL J. Biol. Chem. 265:16576-6581(1990).
 RN [2]
 RP SEQUENCE OF 75-442 FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=YT1;
 RX MEDLINE=88225062; PubMed=3286255;
 RA Kwon S.-T., Terada I., Matsuzawa H., Ohta T.;
 RT "Nucleotide sequence of the gene for aqualysin I (a thermophilic
 alkaline serine protease) of Thermus aquaticus YT-1 and
 characteristics of the deduced primary structure of the enzyme.";
 RL Eur. J. Biochem. 173:491-497(1988).
 RN [3]
 RP SEQUENCE OF 128-170.
 RX MEDLINE=8815937; PubMed=3162211;
 RA Matsuzawa H., Tokugawa K., Hamachi M., Mizoguchi M., Taguchi H.,
 Terada I., Kwon S.-T., Ohta T.;
 RT "Purification and characterization of aqualysin I (a thermophilic
 alkaline serine protease) produced by Thermus aquaticus YT-1.";
 RL Eur. J. Biochem. 171:441-447(1988).
 CC -1- FUNCTION: Aqualysin I is a thermophilic alkaline serine protease.
 The optimal temperature for its caseinolytic activity is 80
 degrees Celsius.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DEVELOPMENTAL STAGE: Secreted from the early stationary phase
 until the time the cells cease to grow.
 CC -1- PTM: THE N- AND C-TERMINAL PRO-SEQUENCES ARE REMOVED THROUGH THE
 PROTEOLYTIC ACTIVITY OF AQUALYSIN I ITSELF, IN THAT ORDER. THE
 C-TERMINAL PRO-SEQUENCE IS REQUIRED FOR TRANSLOCATION OF THE
 PROTEIN ACROSS THE OUTER MEMBRANE.
 CC -1- PTM: Two disulfide bonds are present.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL/ D90108; BAAL135.1; -
 DR EMBL/ X07734; CA30559.1; -
 DR PIR/ A35742; A35742.
 DR HSSP/ P06873; 2PRK.
 DR MEROPS/ S08.051; -
 DR InterPro/ IPR000209; Peptidase_S8.
 DR InterPro/ IPR009020; Protease_inhib.
 DR Pfam/ PF00082; Peptidase_S8; 1.

DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydroxylase/ Serine protease; Zymogen; Signal.
 FT SIGNAL 1
 FT PROPEP 15 127
 FT CHAIN 128 408 AQUALYSIN I.
 FT PROPEP 409 513
 FT ACT SITE 166 166 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 349 349 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 513 AA; 53913 MW; DDFDF6D4A50B785 CRC64;

Query Match 11.6%; Score 259.5; DB 1; Length 513;
 Best Local Similarity 26.4%; Pred. No. 6.6e-10;
 Matches 119; Conservative 45; Mismatches 141; Indels 145; Gaps 24;

QY 16 SYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLY-ALGRTNNANDTNGHTRVAG 74
 DB 152 TYTAGRGVNVYVIDTGIRT-----THREFGRARVGYDALG--GNGQDCNGHGRVAG 203
 QY 75 SYLNGATNGMAQANLVFQSIWD--SSGGLGGLPENTLQTSQAFSAGARTHTN----- 128
 DB 204 TI--GGVTVGAVAVNLVAVRVLDGSGSGTSGVILAGVWV-----TRNHRPVA 252
 QY 129 --SWGAAVNGAYTTDSNVVDYVRKN--DMTILFAAGNEGPNRGRTISAPGTAKNAITVG 183
 DB 253 NMSLGGGVSTA-----LNAVAVNSIAAGVYVAAGNNANAVNS-PARVAELITVG 304
 QY 184 AT--ENLRPSFGSYADINNVAVQSSRGPKDGRKIPDVAPRTYILSAASSLAPDSFV 241
 DB 305 ATTSDDPAPSSNSGCV-----DLFAPGASIPSA-----W 335
 QY 242 AMHDSKAYMGSTMATPIVAGVAVQUREHFVKNRGITP--KPSLIKALLIGAADVGL 298
 DB 336 YTSPTATQTLNGTSMATPHVAG-VALL--YLEQNPATPAPVASAILNGTTRLSIGIS 392
 QY 299 GYPNGNGWGRVTLDKLANVAVNNESSALSTSCAKYTFPATATGKPKLISLVSDAPAST 358
 DB 393 GSPN-----RLLY--SLSSGSGST-----APCTS 415
 QY 359 ---TASVTVLNDLDTLTPNGTRY-----VGNDFSPFNNMDGRN--- 397
 DB 416 CSYITGSLSGGDVNF--QPNRGYIYSPAGTHRAMLRGAGTDFDL-YIMRDGSRMLT 471
 QY 398 -----NVNPFINSPOSGTYTTEVOAYN 420
 DB 472 VGSSTGPTSEBSLSYSTAGYIMRIYAYS 501

RESULT 8
 PROA_VIBAL STANDARD; PRT; 514 AA.
 ID PROA_VIBAL
 AC P16588;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alkaline serine exoprotease A precursor (EC 3.4.21.-).
 GN PROA.
 OS Vibrio alginolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 NCBI_TaxId=663;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89325126; PubMed=2546861;
 RA Deane S.M., Robb F.T., Robb S.M., Woods D.R.;
 RT "Nucleotide sequence of the Vibrio alginolyticus calcium-dependent,
 detergent-resistant alkaline serine exoprotease A.";
 RL Gene 76:281-288(1989).
 CC -1- SIMILARITY: Belongs to peptidase family S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M25499; AAA27550.1; -;
 DR PIR; J50173; J50173.
 DR HSSP; Q99405; IMPT.
 DR MEROPS; S08.050; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 KW Hydrolyase; Serine protease; Zymogen; Signal.
 FT SIGNAL 1 21
 FT PROPEP 22 141
 FT CHAIN 142 534 ALKALINE SERINE EXOPROTEASE A.
 FT ACT_SITE 180 180 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 363 363 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 534 AA; 55930 MW; 84E96D9C649D4226 CRC64;

Query Match 11.2%; Score 252.5; DB 1; Length 534;
 Best Local Similarity 24.3%; Pred. No. 2e-09;
 Matches 116; Conservative 70; Mismatches 137; Indels 155; Gaps 25;

QY 7 IVKADVAQSS--VGL-----YGGQCIYAAVADTGLDTRNNSHEAF 46
 DB 137 IVSADANQTNAIKGLDIDQNLPLDNNYSANFGTGVTVYVDTGV-----NNHVEF 190
 QY 47 RGIITALLY-ALGRTNANDTNGAGTAVAGSVLNGATNKMAPQANLVFQSI--DSGG 103
 DB 191 GGRSVSGYDFVNDADASDCNGHGTVAAGTI---GGLYVAKVNLVGRVLSGSGS 247
 QY 104 LGGPSNLQTLFSQAFAGARIHNSGAVNGAYTTDSNVVDYR---KNDMTLFFA 160
 DB 248 TSGVTAQVWVAANA--SGPSVANSLGGQSYVA-----LDSAVQSAVSGSSMLAA 298
 QY 161 GNEGPNQGTISAPGTAKAATTVGAT--ENDRPSFGSYADININVAQFSSRGPKDRIRK 218
 DB 299 GNSVADACNYS-PARVATGVTVGSTTSDARSSFSMWGSCV----- 338
 QY 219 DYMAPGTYILSARSLAPDSSFMANHSKAYMGSTMATPIYAGNVAQLREHFVKRG 278
 DB 339 DVFPFGSQISLA-----W--YDGYKTIISGSMATPHVAG--VAA--YLGDNSSV 383
 QY 279 TPKPSLKAALIAADAVGLGYPNQGWGRVTLDKSLNAYVNESSALSTOKATYTF 338
 DB 384 S-PSQVEALIVSASR-----GKVT-----DTRGSVVKLXLSLT 416
 QY 339 -----ATACKPL-----KISLWSPAPASTTASVTLV--ND 367
 DB 417 DADGGQCGGDPPTPEDEGLTSGVPSGLSGSGQVAYVYVVEGQRLTVQMYGSGSD 476
 QY 368 LDVITAPNGTRVYGNDFSA-PEDNNMDGR-----NNVENVFINSPOSGTYTIEVCAVN 420
 DB 477 ADVLIR-----FGAKPTLNAMDGRPFYKGNNETCTVSAITSGRHHWIGYS 523

RESULT 9
 ID ALP_CBPAC STANDARD; PRT; 402 AA.
 AC P29118;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alkaline proteinase precursor (EC 3.4.21.-) (ALP).
 GN ALP.
 OS Cephalosporium acremonium (Acremonium chrysogenum).
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Hypocreomycetidae; Hypocreales; Hypocreaceae; mitosporic Hypocreaceae;
 CC Acremonium.
 OK NCBI_TaxID=5044;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=91299283; PubMed=1368696;
 RA Isogai T., Fukagawa M., Kojima H., Kohsaka M., Aoki H., Tanaka H.;
 RT "Cloning and nucleotide sequences of the complementary and genomic
 RL DNA for the alkaline protease from Acremonium chrysogenum.";
 RL Agric. Biol. Chem. 55:471-477(1991).
 CC -1- SIMILARITY: Belongs to peptidase family S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D00923; BAA00765.1; -;
 DR PIR; J00332; J00332.
 DR HSSP; P06873; ZPRK.
 DR MEROPS; S08.DPA; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 KW Hydrolyase; Serine protease; Zymogen; Signal.
 FT SIGNAL 1 20
 FT PROPEP 21 120
 FT CHAIN 121 402 ALKALINE PROTEINASE.
 FT ACT_SITE 160 160 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 191 191 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 347 347 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 402 AA; 42099 MW; BD030CCD42D918E1 CRC64;

Query Match 11.2%; Score 250.5; DB 1; Length 402;
 Best Local Similarity 31.0%; Pred. No. 1.9e-09;
 Matches 93; Conservative 31; Mismatches 101; Indels 75; Gaps 14;

QY 21 GGGQIVAVADTGLTGRNDSMEAFRGK-ITALIYALGRTNANDTNGHGTVAAGSVLGN 79
 DB 151 GSGTYAAVVDVGT-----LESHNEFSGRAITGNVAGGSEN--ADTNGHGTVAAGTI--- 199
 QY 80 GATNKMAPQANLVFQSIYDSSG-----LGGLPSNLTQFSQAFSAQARI-----HT 127
 DB 200 GGRITGVAKNNLNAVYFRSSSSSTSIILDFPWAIVNDIINRQRKALISLGGGYS 259
 QY 128 NSWGAAYNGAYTTSSKAVVDYRKNDMTILFAAGNEPNQGTISAPGTAKAATTVG--AT 185
 DB 260 SAFNNAVNTAY--SRGLSVY-----AANDNQANYS-PASANAATIVGSIAS 306
 QY 186 ENLRPSFGSYADININVAQFSSRGPTDGRIRKPYMAPGTYILSARSLAPDSSFMANHD 245
 DB 307 NMASSSFNYSGLV-----DIFAPGISLSA-----WIGGN 337
 QY 246 SKIAYMGSTMAATPIYAGNVAQLREHFVKRGITPKPSLKAALIAADAV--GLGYEN 302
 DB 338 SATNTISGTSWATPHVTVGLYDQ-----ALEGLTSGAAARLINALATTGTVSNFGSGSPN 393

RESULT 10
 ID WPA_BACSU STANDARD; PRT; 894 AA.
 WPA_BACSU

CC P54423; 006726;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell
 wall-associated polypeptides CWBP23 and CWBP52].
 GN WPRA OR BSU10770
 OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=1423;
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
 RC STRAIN=168;
 RC MEDLINE=97158234; PubMed=9004506;
 RT Margot P., Karamata D.;
 RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
 growth, encodes a cell-wall-associated protease.";
 RL Microbiology 142:3437-3444(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RC MEDLINE=98015415; PubMed=9353931;
 RT Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.,
 RT Sequencing of regions downstream of addA (98 degrees) and citG (289
 degrees) in Bacillus subtilis.";
 RL Microbiology 143:3305-3308(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RC MEDLINE=98044033; PubMed=9384377;
 RT Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RT Aaeved V., Bertero M.G., Bessieres P., Bolojin A., Borcherdt S.,
 RT Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RT Brouillet S., Bruschi C.V., Caldwell B., Capriano V., Carter N.M.,
 RT Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RT Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RT Ertlen K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RT Fritz C., Fujita M., Fujita Y., Funa S., Gallizi A., Galleron N.,
 RT Gilm S.Y., Glaeser P., Goffeau A., Goldschly E.J., Grandi G.,
 RT Galispi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RT Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RT Jorib B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RT Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RT Kurita K., Lapidus A., Lardinois S., Rauber J., Lazarevic V.,
 RT Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medgye C.,
 RT Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RT Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
 RT Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RT Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RT Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RT Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RT Sekiuchi U., Sekowska A., Seror S.J., Seror P., Shin B.S., Solido B.,
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Toso V., Uchiyama S., Vandenbol M., Vannier F., Vasseroletti A.,
 RA Varsi A., Wambit R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zuretein E., Yoshikawa H., Darchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis".
 RT Nature 390:249-256(1997).
 RL
 CC - FUNCTION: NOT YET KNOWN; COULD BE INVOLVED IN PROTEOLYTIC
 CC DEGRADATION, BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
 CC - SUBCELLULAR LOCATION: Cell-wall bound.
 CC - PTM: PROCESSED INTO CWBP23 AND CWBP52.
 CC - SIMILARITY: Belongs to peptidase family S8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL; U58981; AAC25926.1; -
 CC EMBL; Y09476; CA170641.1; -
 CC EMBL; Z99109; CAB12917.1; -
 CC F01; F69730; F69730.
 CC HSP; Q45670; DBI.
 CC MER05; S08.004; -
 CC Subtilisin; Bg1846; WPRA.
 CC Interpro; IPR00209; Peptidase_S8.
 CC Pfam; PF00082; Peptidase_S8; 1.
 CC PRINTS; PRO0723; SUBTILISIN.
 CC PROSITE; PS00136; SUBTILISIN ASP; FALSE_NEG.
 CC PROSITE; PS00137; SUBTILISIN HIS; 1.
 CC PROSITE; PS00138; SUBTILISIN SER; 1.
 CC Hydrolyase; Serine protease; Cell wall; Zymogen; Signal;
 CC Complete proteome.
 CC SIGNAL 1 31
 CC FT CHAIN 32 894 CELL WALL-ASSOCIATED PROTEASE.
 CC FT CHAIN 32 894 CWBP23.
 CC FT PROPEP 2 413 POTENTIAL.
 CC FT CHAIN 414 894 CWBP52.
 CC FT ACT_SITE 466 466 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 650 650 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT CONFLICT 9 9 V -> A (IN REF. 1).
 CC FT CONFLICT 14 14 L -> I (IN REF. 1).
 CC SQ SEQUENCE 894 AA; 96487 MW; 0F67C353E5F8DBC CRC64;
 CC
 CC Query Match 11.18; Score 249.5; DB 1; Length 894;
 CC Best Local Similarity 24.28; Pred. No. 6.1e-09;
 CC Matches 106; Conservative 63; Mismatches 140; Indels 129; Gaps 19;
 CC
 CC 25 IVAAVDGLTGRNDSMEHAFKGTALVA---LGRNNANDTNGHGTAVAGSYG--- 78
 CC 457 LIAVVDGVSTLAD-----LKKVRTDIGHFVGRNNAMMDQGHGTVAGIIAASD 510
 CC 79 NGATNKMAFQANLVFQSIWSSGGIQLPSNLQTLFSQAFSAGARLHTNSWGAANGAY 138
 CC 511 NGYMTGLNKKAKIIPKVLDSAG--SGDEQIALGIKYAADGAKVIMSLG---GGV 564
 CC 139 TTDSRNVDDVVRKNMILFAAGNEGNGTISAPGAKAIVGATENLRPSFGSYAON 198
 CC 565 SVLEPLKTKAADKRVILAAAGNDEN--ALSYPSASKTVMSVGT-----NR 611
 CC 199 INHVAQSSRGPTDGRKEDVVAFGTYILSARSSLAPOSFPAANDSKXAYVGTSMAT 258
 CC 612 MDMTADPSNYGKL-----DISAPGSDI---PSLVPNGN-----VTYMSGTSMAT 653
 CC 259 PIVAGNVAGLRHENVKRGI--TPKPSLLAALIAAGADVGLGYPNGQ----- 305
 CC 654 PYAAAAAGLL--FAQPKTKTEVEDMLKKT---ADTSFSSVGGEBELDYDDP 706
 CC 306 -----GQRYTLDKSLNVA---YNESSALSTSQAT-----Y 335
 CC 707 EIKPTGVDMHSGGRNWKAVASADLQKVKKLSSTQYAVGSAKEGTLEWNGKK 766
 CC 336 TETITAK--PLKISLYMSAPASTASVTLVNDLDLVITAPNGTRVNGDFAPDNW 353
 CC 767 LGSAAKAKDNANFKNI-----ATQKODQVLYLKATRG----- 798
 CC
 CC QY 394 DGRNNVNFINSQSGT 411
 CC DB 799 DAKTSYVVVVKKGPST 816
 CC
 CC RESULT 11
 CC SUBS_BACLE STANDARD; PRT; 269 AA.
 CC ID SUBS_BACLE
 CC AC P29600;
 CC DT 01-APR-1993 (Rel. 25, Created)
 CC DT 01-APR-1993 (Rel. 25, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Subtilisin Savinase (EC 3.4.21.62) (Alkaline protease).
 OS Bacillus lentus.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1467;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RX MEDLINE=92148829; PubMed=1738156;
 RA Betzel C., Klupsch S., Papendorf G., Hastrup S., Branner S.,
 Wilson K.S.;
 RT "Crystal structure of the alkaline proteinase Savinase from Bacillus
 lentus at 1.4-A resolution".
 RL J. Mol. Biol. 223:427-445(1992).
 RN [2]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96184541; PubMed=8654411;
 RA Remerowski M.L., Pepermans H.A.M., Hilbers C.W., van de Ven F.J.M.;
 RT "Backbone dynamics of the 269-residue protease Savinase determined
 from 15N-NMR relaxation measurements".
 RL Eur. J. Biochem. 235:629-640(1996).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (0.78 ANGSTROMS).
 RX MEDLINE=98426039; PubMed=9753430;
 RA Kuhn P., Knapp M., Soltis S.M., Ganshaw G., Thoenes M., Bott R.;
 RT "The 0.78-A structure of a serine protease: Bacillus lentus
 subtilisin".
 RL Biochemistry 37:13446-13452(1998).
 CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 it catalyzes the hydrolysis of proteins and peptide amides.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 for peptide bonds, and a preference for a large uncharged residue
 in P1. Hydrolyzes peptide amides.
 CC -1- COFACTOR: Binds 2 calcium ions per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
 Savinase by Novozymes.
 CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 sporulation, and many mutations which block sporulation at early
 stages affect expression levels of subtilisin. However, subtilisin
 is not necessary for normal sporulation.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 DR PDB: 1C9J; 06-OCT-99.
 DR PDB: 1C9N; 10-JAN-01.
 DR PDB: 1C9M; 10-JAN-01.
 DR PDB: 1GCI; 11-NOV-98.
 DR PDB: 1IAV; 11-JUL-01.
 DR PDB: 1JEA; 26-NOV-97.
 DR PDB: 1SVN; 14-OCT-96.
 DR MEROPS: S08.103.
 DR InterPro: IPR000203; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE ASP; 1.
 DR PROSITE: PS00137; SUBTILASE HIS; 1.
 DR PROSITE: PS00138; SUBTILASE SER; 1.
 KW Hydrolyase; Sporulation; Serine protease; Metal-binding;
 KM Calcium-binding; 3d-structure
 FT ACT_SITE 32 32 CHANGE RELAY SYSTEM.
 FT ACT_SITE 62 62 CHANGE RELAY SYSTEM.
 FT ACT_SITE 215 215 CHANGE RELAY SYSTEM.
 FT METAL 2 2 CALCIUM 1.
 FT METAL 40 40 CALCIUM 1.
 FT METAL 73 73 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 75 75 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 77 77 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 79 79 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 163 163 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 165 165 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 168 168 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT STRAND 2 2
 FT HELIX 6 6
 FT TURN 11 10
 FT TURN 12 12
 FT HELIX 13 18
 FT HELIX 19 20
 FT TURN

FT TURN 24 25
 FT STRAND 27 32
 FT TURN 39 40
 FT STRAND 43 48
 FT TURN 51 52
 FT HELIX 62 71
 FT STRAND 78 78
 FT TURN 84 85
 FT STRAND 87 92
 FT STRAND 94 94
 FT TURN 96 97
 FT STRAND 100 100
 FT HELIX 102 114
 FT TURN 115 116
 FT STRAND 119 122
 FT STRAND 126 126
 FT HELIX 131 142
 FT TURN 143 144
 FT STRAND 146 150
 FT STRAND 161 161
 FT TURN 162 164
 FT TURN 166 167
 FT STRAND 168 174
 FT TURN 176 177
 FT STRAND 180 180
 FT TURN 182 183
 FT TURN 188 189
 FT STRAND 190 195
 FT STRAND 199 203
 FT TURN 204 206
 FT STRAND 207 211
 FT HELIX 214 231
 FT TURN 233 234
 FT TURN 237 246
 FT TURN 247 247
 FT STRAND 249 249
 FT HELIX 254 257
 FT TURN 258 259
 FT STRAND 261 261
 FT HELIX 264 267
 FT TURN 268 269
 SQ SEQUENCE 269 AA; 26698 MW; 4D89F8778999BFD CRC64;
 Query Match 11.0%; Score 246; DB 1; Length 269;
 Best Local Similarity 31.4%; Pred. No. 2,28-09;
 Matches 92; Conservative 30; Mismatches 89; Indels 82; Gaps 14;
 QY 8 VKADYAGSSYGLYGGQIVAVADTGLDTGRNDSMEAFKRTALYALGRTNANDTNG 67
 DB 11 VOAPAAHNR-GLTSGGVAVAVLDTGIST-----HPDLNIRGASFPSP--STODGNG 61
 QY 68 HGTHVAGSV--LGNGATNKGAPOANLVFGSIMDSGGLGSPENLQTLFSQAFSAPARI 125
 DB 62 HGTHVAGTIALNNSIGVLGVAIPSAELYAVVYLAAG--SGSVSIAQGLEMGANMNV 119
 QY 126 HTNWKGA-----AVNGAYTTDSRNVDDYKRNMTLLFAAGNCGPNGGTISAFGTA 176
 DB 120 ANLSIGSPSPATLEQAVNSA---TSRGV-----LVVAASGNNG--AGSISYPARY 165
 QY 177 KQAIIVAGTE--NRPSPGSYADNINHVAPSSSGPTKDGIKPDVWAPGYIILASAPSL 234
 DB 166 ANNAVAGTIDQNNRNASISQYAGI-----DIVAFGVNVOSTYTPG- 205
 QY 235 APDSGFMANHDSKYAYMGTSMAPIYAGNYA-----QUREHFVKN 275
 DB 206 -----STYASLNGTSMATPHVAGAAALVYKQKPSMSNVQIRNH-LKN 246
 RESULT 12
 ELYA_BACAO STANDARD; PRT; 380 AA.
 ID ELYA_BACAO
 AC P27693;
 DT 01-AUG-1992 (Rel. 23, Created)

DT 10-OCT-2003 (Rel. 42, last annotation update)
 DE Alkaline protease precursor (EC 3.4.21.-).
 OS *Bacillus clausii*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=9880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
 RX MEDLINE=93043753; PubMed=1368952;
 RA Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,
 RA Aono R., Horikoshi K.,
 RT "Molecular cloning, nucleotide sequence, and expression of the
 RT structural gene for alkaline serine protease from alkaliphilic
 RT *Bacillus* sp. 221."
 RL Biosci. Biotechnol. Biochem. 56:1455-1460(1992).
 RP [2]
 RC SEQUENCE OF 112-129.
 RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
 RA Horikoshi K.;
 RL (in) Horikoshi K. (eds.);
 RL Microorganisms in alkaline environments, pp.187-194, VCH,
 RL Weinheim (1991).
 CC -1- COFACTOR: Binds 2 calcium ions per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; S46754; AAC60420.1; -
 DR EMBL; D13157; BAA02442.1; -
 DR EMBL; A26817; CAA01836.1; -
 DR EMBL; A22550; CAA01611.1; -
 DR HSP; P29600; IGC1.
 DR MEROPS; S08.103; -
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Hydrolyase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
 KM Signal.
 KM Signal.
 FT SIGNAL 1 27
 FT PROPEP 28 111
 FT CHAIN 112 380
 FT ACT_SITE 143 143
 FT ACT_SITE 173 173
 FT ACT_SITE 326 326
 FT METAL 113 113
 FT METAL 151 151
 FT METAL 184 184
 FT METAL 186 186
 FT METAL 188 188
 FT METAL 190 190
 FT METAL 274 274
 FT METAL 276 276
 FT METAL 279 279
 FT SEQUENCE 380 AA; 38826 MW; 5F73ABCG8DB6831 CRC64;
 Query Match 11.0%; Score 246; DB 1; Length 380;

Best Local Similarity 31.4%; Pred. No. 3.4e-09;
 Matches 92; Conservative 30; Mismatches 89; Indels 82; Gaps 14;
 QY 8 VKADVAGSSVGLYQGGIIVAVADTGLDTGRNDSMEHAFRGKITAYALGRTNANDTNG 67
 DB 122 VQAPAAHNR-GLTNGSGYVAVLTDTGIST-----HFDLNIIRGASFPGEF-STQDGG 172
 QY 68 HGTAVAGSV--LGRGATNKGAPOANLVFQSIINDSGGLGEPNLDTLTSQAFSAGARI 125
 DB 173 HGTAVAGTIALNNSTIGLVGAPSAELVAYKVLGASG--SGVSSIQGLEMGAGNMGHV 230
 QY 126 HTNSMGA-----AVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGFGTISAPGTA 176
 DB 231 ANLSIGSPSPATLEQAVNSA---TSRGV-----LVVAASNGG--AGSISYPARY 276
 QY 177 KNATVAGTE--NRPSFGSYADININVAFPSSRGPTKDRIRKPDVWAPGTIILSARSL 234
 DB 277 ANMAVAGHTDQNNRPSFSQIGAGL-----DIVAPGVNVQSTYPG- 316
 QY 235 APDSFMANHDSKYAVNGTSMATPIVAGNVA-----QLREHFVKV 275
 DB 317 STVASLNGTSMATPIVAGAAALVKQNPFSMNVQIRNH-LKN 357
 RESULT 14
 ID THES_BACSP STANDARD; PRT; 401 AA.
 AC Q45670;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 10-OCT-2003 (Rel. 42, last annotation update)
 DE Thermophilic serine proteinase precursor (EC 3.4.21.-) (Ak.1
 DE protease).
 OS *Bacillus* sp. (strain AK1).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1409;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9508562; PubMed=7993087;
 RA Maciver B., McHale R.H., Saul D.J., Bergquist P.L.;
 RT "Cloning and sequencing of a serine proteinase gene from a
 RT thermophilic *Bacillus* species and its expression in *Escherichia*
 RT *coli*."
 RL Appl. Environ. Microbiol. 60:3981-3988(1994).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
 RX MEDLINE=20057863; PubMed=1058904;
 RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
 RT "Calcium-mediated thermostability in the subtilisin superfamily: the
 RT crystal structure of *Bacillus* Ak.1 protease at 1.8-A resolution";
 RL J. Mol. Biol. 294:1027-1040(1999).
 CC -1- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: Has a pH optimum of 8.5, a temperature optimum of
 CC 75 degrees Celsius.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; L29506; AAA63688.1; -
 DR PIR; J39974; J39974.
 DR PDB; 1DBT; 18-NOV-99.
 DR MEROPS; S08.009; -
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 KW Hydrolyase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
 KW Signal; 3D-structure.
 FT SIGNAL 1 24
 FT PROPEP 25 121
 FT CHAIN 122 401
 FT ACT_SITE 160 160
 FT ACT_SITE 193 193
 FT ACT_SITE 347 347
 FT METAL 126 126
 FT METAL 168 168
 FT METAL 169 169
 FT METAL 171 171
 FT METAL 179 179
 FT METAL 184 184
 FT METAL 186 186
 FT METAL 204 204
 FT METAL 204 204
 FT METAL 207 207
 FT METAL 209 209
 FT METAL 211 211
 FT METAL 297 297
 FT METAL 300 300
 FT METAL 323 323
 FT DISULFID 258 260
 FT TURN 127 128
 FT HELIX 129 132
 FT TURN 135 136
 FT HELIX 137 139
 FT TURN 140 140
 FT HELIX 141 144
 FT TURN 145 147
 FT TURN 152 153
 FT STRAND 155 160
 FT TURN 165 166
 FT TURN 168 173
 FT STRAND 174 179
 FT TURN 180 183
 FT STRAND 184 184
 FT HELIX 193 202
 FT STRAND 219 224
 FT TURN 228 229
 FT HELIX 234 246
 FT TURN 247 248
 FT STRAND 251 254
 FT HELIX 263 274
 FT TURN 275 276
 FT STRAND 278 282
 FT STRAND 285 285
 FT TURN 298 299
 FT STRAND 301 306
 FT TURN 308 309
 FT STRAND 312 312
 FT TURN 314 315
 FT STRAND 316 316
 FT TURN 320 321
 FT STRAND 324 327
 FT STRAND 331 335
 FT TURN 336 338
 FT STRAND 339 343
 FT HELIX 346 364
 FT TURN 363 364
 FT HELIX 367 376
 FT TURN 377 377
 FT STRAND 379 379
 FT TURN 382 383
 FT STRAND 384 384
 FT TURN 385 387
 FT STRAND 388 388
 FT STRAND 391 392
 FT HELIX 395 399

POTENTIAL.
 THERMOPHILIC SERINE PROTEINASE.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 CALCULUM 1.
 CALCULUM 3 (VIA CARBONYL OXYGEN).
 CALCULUM 1.
 CALCULUM 1.
 CALCULUM 2.
 CALCULUM 2 (VIA CARBONYL OXYGEN).
 CALCULUM 1.
 CALCULUM 3.
 CALCULUM 1.
 CALCULUM 1 (VIA CARBONYL OXYGEN).
 CALCULUM 1 (VIA CARBONYL OXYGEN).
 SODIUM (VIA CARBONYL OXYGEN).
 SODIUM (VIA CARBONYL OXYGEN).
 SODIUM.

FT TURN 400 401
 SQ SEQUENCE 401 AA; 42835 MW; 1C736EFA89F256CF CRC64;
 Query Match 10.9%; Score 245.5; DB 1; Length 401;
 Best Local Similarity 31.7%; Pred. No. 3.9e-09;
 Matches 83; Conservative 29; Mismatches 99; Indels 51; Gaps 11;
 QY 11 DVAOSYGLYQGGIVAVADTGLDGRNDSMEHAFGKITALY-ALGRTNADVTGHH 69
 DB 146 DVTGSS-----SGEIAVIDGV-----YTHDDLQKVIKGFDPNDNDYDPLNHHG 194
 QY 70 THVAGSVIG-NGATN-KGAPQANLYFQGISDSSGGLGSPNLQTLFQAPSAGARIH 126
 DB 195 THVAGIAAETNNATGAGAPNTRILAVRALDENG--SGTLSDIADAITVAADSGAEVI 252
 QY 127 TNSWGAANGAVTDSRVNDYVRKNTMTLFAAGNGEPNGGTISAPGTAKNAITYGATE 166
 DB 253 NLSLGC--DHTTTLNNAVYANKGSVVYAAAGNNG--SSTTFEPASTENYIAVGA-- 305
 QY 187 NLRPSFGSYADNINHVAFQSSRGPTKDRIKPDYMAEGTYILSARSSLAPDSSFWANHS 246
 DB 306 -----VDQYRLASFVNYGTW-----VDVAVGVDIVSHITG-----N 338
 QY 247 KYAVMGTSMTPTVAGVNAQL 268
 DB 339 KYAVMGTSMTASPHVAGLAALL 360
 RESULT 15
 PRM_BACSP STANDARD; PRT; 269 AA.
 ID PRM_BACSP
 AC Q99405
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DD M-protease (EC 3.4.21.-)
 OS Bacillus sp. (strain KSM-K16).
 OC Bacteria; Firmicutes; Bacillales; Bacillus.
 NX NCB Taxid=1409;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RA Yamane T., Kani T., Hatanaka T., Suzuki A., Ashida T.,
 RA Kobayashi T., Ito S., Yamashita O.,
 RT "Structure of a new alkaline serine protease (M-protease) from
 RT Bacillus sp. KSM-K16".
 RL Acta Crystallogr. D 51:199-206 (1995).
 RN [2]
 RP SEQUENCE OF 1-23, AND CHARACTERIZATION.
 RX MEDLINE=9535832; PubMed=7632397;
 RA Kobayashi T., Hakamada Y., Adachi S., Hitomi J., Yoshimatsu T.,
 RA Koike K., Kawai S., Ito S.;
 RT "Purification and properties of an alkaline protease from
 RT alkalophilic Bacillus sp. KSM-K16".
 RL Appl. Microbiol. Biotechnol. 43:473-481 (1995).
 CC -1- COFACTOR: Binds 2 calcium ions per subunit.
 CC -1- SUBUNIT: Monomer.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 DR PDB: 1MPT; 22-JUN-94.
 DR MEROPS: S08.010; -;
 DR InterPro: IPR0000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 KW Hydrolyase; Serine protease; Metal-binding; Calcium-binding;
 KW 3D-structure.
 FT ACT_SITE 32 32
 FT ACT_SITE 62 62
 FT ACT_SITE 215 215
 FT METAL 2 2
 FT METAL 40 40
 FT METAL 73 73
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 CALCULUM 1.
 CALCULUM 1.
 CALCULUM 1 (VIA CARBONYL OXYGEN).

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 32.55 Seconds
(without alignments)
4206.909 Million cell updates/sec

Title: US-09-985-689a-2

Perfect score: 2245

Sequence: 1 NDVARGIVKADVAQSSYGVY.....EVQAVNPVGPQNSIAIVN 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2237	99.6	639	2 Q9AOR3	Q9AGR3 bacillus sp
2	2191	97.6	640	2 Q9AOR3	Q9AUV9 bacillus sp
3	2130	94.9	434	2 Q9AOR0	Q9AGR0 bacillus sp
4	2010.5	89.6	433	2 Q9AOR1	Q9AGR1 bacillus sp
5	2006.5	89.4	433	2 Q9AOR4	Q9AGR4 bacillus sp
6	1999.5	89.1	433	2 Q9AOR2	Q9AGR2 bacillus sp
7	509.5	22.7	1925	5 Q8TSM1	Q8TSM1 dictyosteli
8	443	19.7	1702	5 Q9GTN7	Q9GTN7 dictyosteli
9	423	18.8	654	17 Q8UC09	Q8UC09 pyrococcus
10	418	18.6	561	16 Q8RBU2	Q8RBU2 thermococcus
11	407.5	18.2	1239	16 Q9PB24	Q9PB24 streptomyc
12	387	17.2	430	16 Q8ENV1	Q8ENV1 oceanobacil
13	385	17.1	1253	16 Q9FC06	Q9FC06 streptomyc
14	365	16.3	1102	2 P95684	P95684 streptomyc
15	360.5	16.1	1208	16 Q82B14	Q82B14 streptomyc
16	357.5	15.9	444	16 Q9KBU7	Q9KBU7 bacillus ha

17	354.5	15.8	1105	2 Q8KX6	Q8KX6 streptomyc
18	351	15.6	1245	16 Q8L54	Q8L54 streptomyc
19	345	15.4	1139	16 Q8L139	Q8L139 streptomyc
20	343	15.3	1237	2 Q8GGT4	Q8GGT4 streptomyc
21	340	15.1	412	2 Q9AER6	Q9AER6 thermococcus
22	339	15.1	412	16 Q9AER6	Q9AER6 thermococcus
23	331.5	14.8	435	16 Q8EMJ3	Q8EMJ3 oceanobacil
24	329.5	14.7	1220	16 Q9L0A0	Q9L0A0 streptomyc
25	326	14.5	442	16 Q8J788	Q8J788 bacillus su
26	323.5	14.4	851	1 Q93635	Q93635 thermococcus
27	315.5	14.1	824	2 Q45464	Q45464 bacillus sp
28	306.5	13.7	1398	1 Q9P9L1	Q9P9L1 pyrococcus
29	301.5	13.4	431	2 Q9S3I6	Q9S3I6 bacillus sp
30	289	12.9	1135	1 Q9P9D1	Q9P9D1 uncultured
31	287.5	12.8	434	2 Q54327	Q54327 bacillus sp
32	283	12.6	419	2 Q45681	Q45681 bacillus su
33	280.5	12.5	586	16 Q8PAL8	Q8PAL8 xanthomonas
34	279	12.4	799	16 Q9KEM1	Q9KEM1 bacillus ha
35	275.5	12.3	959	16 Q8PMS7	Q8PMS7 xanthomonas
36	274.5	12.2	1345	1 Q54437	Q54437 staphylothe
37	274	12.2	1098	16 Q9L1Z8	Q9L1Z8 streptomyc
38	272.5	12.1	621	2 Q9F486	Q9F486 alteromonas
39	272.5	12.1	621	2 Q53401	Q53401 alteromonas
40	270	12.0	715	2 P70765	P70765 alteromonas
41	270	12.0	1101	16 Q82CFO	Q82CFO streptomyc
42	269.5	12.0	568	16 Q8PMC0	Q8PMC0 xanthomonas
43	269.5	12.0	1571	2 Q8GCM3	Q8GCM3 streptococc
44	268.5	12.0	629	2 Q93R88	Q93R88 alteromonas
45	265.5	11.8	403	2 Q45463	Q45463 bacillus sp

ALIGNMENTS

RESULT 1

Q9AOR3 PRELIMINARY; PRT; 639 AA.
ID Q9AOR3
AC Q9AOR3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease.
CN PROA.
OS Bacillus sp. 9860.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_taxid=133778;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9860;
RX MEDLINE=20568675; PubMed=1118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
Horikoshi K.
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046403; BAB31266.2; -.
DR HSSP; P00782; 1SUP.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_SER; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 639 AA; 68185 MW; 316AR6FPDB24FF54 CRC64;

Query Match 99.6%; Score 2237; DB 2; Length 639;
Best Local Similarity 99.8%; Pred. No. 5.2e-121;
Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAOSSYGLGQGIIVAVADTGLDGTGRNDSMHEAFRGKITLALYALGRN 60
 DB 206 NDVARGIVKADVAOSSYGLGQGIIVAVADTGLDGTGRNDSMHEAFRGKITLALYALGRN 265
 QY 61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 120
 DB 266 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 325
 QY 121 AGARHTNSMGAANNGVYTTDSRVDDYRKNDMTLLFAAGNGPVGGTISAGTAKNAI 180
 DB 326 AGARHTNSMGAANNGVYTTDSRVDDYRKNDMTLLFAAGNGPVGGTISAGTAKNAI 385
 QY 181 TVGATEMLRPSFGSYADINNHVAQFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 240
 DB 386 TVGATEMLRPSFGSYADINNHVAQFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 445
 QY 241 WANHDSKYAAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAAGADVGLG 300
 DB 446 WANHDSKYAAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAAGADVGLG 505
 QY 301 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
 DB 506 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 565
 QY 361 SVTLVNDLDVITAPNGTRVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 420
 DB 566 SVTLVNDLDVITAPNGTRVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 625
 QY 421 VPVGPQNFSLAIYN 434
 DB 626 VPVGPQNFSLAIYN 639

RESULT 2
 Q93UV9 PRELIMINARY; PRT; 640 AA.
 AC Q93UV9;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Protease.
 GN PROF.
 OS *Bacillus* sp. KSM-KP43.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=109322;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KP43;
 RA Itoh S., Saeki K.,
 RT "new protease."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB051423; BAB55674.2; -
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002039; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 SQ SEQUENCE 640 AA; 67991 MW; 48BAF79D592C15 CRC64;

Query Match 97.6%; Score 2191; DB 2; Length 640;
 Best Local Similarity 96.5%; Pred. No. 2,3e-118;
 Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAOSSYGLGQGIIVAVADTGLDGTGRNDSMHEAFRGKITLALYALGRN 60
 DB 207 NDVARGIVKADVAOSSYGLGQGIIVAVADTGLDGTGRNDSMHEAFRGKITLALYALGRN 266
 QY 61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 120

DB 267 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 326
 QY 121 AGARHTNSMGAANNGVYTTDSRVDDYRKNDMTLLFAAGNGPVGGTISAGTAKNAI 180
 DB 327 AGARHTNSMGAANNGVYTTDSRVDDYRKNDMTLLFAAGNGPVGGTISAGTAKNAI 386
 QY 181 TVGATEMLRPSFGSYADINNHVAQFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 240
 DB 387 TVGATEMLRPSFGSYADINNHVAQFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 446
 QY 241 WANHDSKYAAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAAGADVGLG 300
 DB 447 WANHDSKYAAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAAGADVGLG 506
 QY 301 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
 DB 507 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 566
 QY 361 SVTLVNDLDVITAPNGTRVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 420
 DB 567 SVTLVNDLDVITAPNGTRVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 626
 QY 421 VPVGPQNFSLAIYN 434
 DB 627 VPVGPQNFSLAIYN 640

RESULT 3
 Q9AOR0 PRELIMINARY; PRT; 434 AA.
 AC Q9AOR0;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROB.
 OS *Bacillus* sp. NVL.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=133781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NVL;
 RA MEDLINE=20568675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RT Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 alkaliphilic *Bacillus* spp.: enzymatic properties, sequences, and
 evolutionary relationships."
 RT Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -1-SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046406; BAB21269.1; -
 DR HSSP; P00782; ISUP.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002039; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;

Query Match 94.9%; Score 2130; DB 2; Length 434;
 Best Local Similarity 93.4%; Pred. No. 4.6e-115;
 Matches 401; Conservative 24; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAOSSYGLGQGIIVAVADTGLDGTGRNDSMHEAFRGKITLALYALGRN 60

Db 1 NDVARGIVKADVAOSSGYLGOGVVAVADTGLDTRGNDSSMHEARFGKITAIYALGRTN 60
 QY 61 NANDTNGHGTTHVAGSYLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTQTLFSGQAFS 120
 Db 61 NANDPNGHGTTHVAGSYLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTQTLFSGQAFS 120
 QY 121 AGARIHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGGTISAPGTAKNAI 180
 Db 121 AGARIHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGGTISAPGTAKNAI 180
 QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAFGTYILSARSILAPDSSF 240
 Db 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAFGTYILSARSILAPDSSF 240
 QY 241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
 Db 241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
 QY 301 PNGNOGGRVTLDKSLNVAAYVNESSALSTSQKATYTFRTAGKPLKISLWSDAPASTTA 360
 Db 301 PNGNOGGRVTLDKSLNVAAYVNESSALSTSQKATYTFRTAGKPLKISLWSDAPASTTA 360
 QY 361 SYTLVNDLVLVITAPNGTRYVGNDFSAFPDNNMDGNNEVNFINSPOGTYTIEVOAYN 420
 Db 361 SYTLVNDLVLVITAPNGTRYVGNDFSAFPDNNMDGNNEVNFINSPOGTYTIEVOAYN 420
 QY 421 VPVGPQNFSLATYN 434
 Db 421 VPVGPQNFSLATYN 434

RESULT 4

Q9AQR1 ID Q9AQR1 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR1; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. SD521.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NC NCB1_TaxID=133780;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SD521;
 RX MEDLINE=2056675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H., Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from alkaliphilic Bacillus spp.: enzymatic properties, sequences, and evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046405; BAB21268.1; -.
 DR HSSP; Q45670; IDB1.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002029; Peptidase_S8.
 DR InterPro; IPR007280; Peptidase_S8.
 DR Pfam; PF04151; PFC; 1.
 DR Pfam; PF04151; PFC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 FT NON_TER 1 1
 FT 433 433
 SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18F660DC CRC64;

Query Match 89.6%; Score 2010.5; DB 2; Length 433;
 Best Local Similarity 88.5%; Pred. No. 3.5e-108;

Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAOSSGYLGOGVVAVADTGLDTRGNDSSMHEARFGKITAIYALGRTN 60
 Db 1 NDVARGIVKADVAOSSGYLGOGVVAVADTGLDTRGNDSSMHEARFGKITAIYALGRTN 60
 QY 61 NANDTNGHGTTHVAGSYLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTQTLFSGQAFS 120
 Db 61 NANDPNGHGTTHVAGSYLNGATNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTQTLFSGQAFS 119
 QY 121 AGARIHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGGTISAPGTAKNAI 180
 Db 120 AGARIHTNSWGAAYNGAYTNRSQVDEYVRKNDMTILFAAGNEGPGGTISAPGTAKNAI 179
 QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAFGTYILSARSILAPDSSF 240
 Db 180 TVGATENYRPSFSLADNPNHIAQFSSRGATRDGRIKPDVTAAGTYILSARSILAPDSSF 239
 QY 241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
 Db 240 WANHNSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 299
 QY 301 PNGNOGGRVTLDKSLNVAAYVNESSALSTSQKATYTFRTAGKPLKISLWSDAPASTTA 360
 Db 300 PNGNOGGRVTLDKSLNVAAYNEATATATGKATYFQAQAGKPLKISLWSDAPASTTA 359
 QY 361 SYTLVNDLVLVITAPNGTRYVGNDFSAFPDNNMDGNNEVNFINSPOGTYTIEVOAYN 420
 Db 360 SYTLVNDLVLVITAPNOKYVGNDFSYPDNNMDGNNEVNFINSPOGTYTIEVOAYN 419
 QY 421 VPVGPQNFSLATYN 434
 Db 420 VPVGPQNFSLATYN 433

RESULT 5

Q9AQR4 ID Q9AQR4 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR4; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. D6.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NC NCB1_TaxID=127889;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D6;
 RX MEDLINE=2056675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H., Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from alkaliphilic Bacillus spp.: enzymatic properties, sequences, and evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046402; BAB21265.1; -.
 DR HSSP; Q45670; IDB1.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002029; Peptidase_S8.
 DR InterPro; IPR007280; Peptidase_S8.
 DR Pfam; PF04151; PFC; 1.
 DR Pfam; PF04151; PFC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 FT NON_TER 1 1
 FT 433 433

SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;

Query Match 89.4%; Score 2006.5; DB 2; Length 433;
 Best Local Similarity 88.2%; Pred. No. 6e-108;
 Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGIDTGRNDSNHEAFRGKITLALYALGRTN 60
 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGIDTGRNDSNHEAFRGKITLALYALGRTN 60

DB 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGIDTGRNDSNHEAFRGKITLALYALGRTN 60

QY 61 NANDINGHGHVAGSVLNGATNKGAPQANLVFQSIMSSGGLGGLPENTLTFEQAN 120
 61 NANDINGHGHVAGSVLNGATNKGAPQANLVFQSIMSSGGLGGLPENTLTFEQAN 120

DB 61 NANDINGHGHVAGSVLNGATNKGAPQANLVFQSIMSSGGLGGLPENTLTFEQAN 120

QY 121 AGARHTNSMGAAYNGAYTTDSRNVDYRKNDMTLFAAGNEPNGGTTISAPGTAKNAI 180
 121 AGARHTNSMGAAYNGAYTTDSRNVDYRKNDMTLFAAGNEPNGGTTISAPGTAKNAI 180

DB 120 AGARHTNSMGAAYNGAYTTDSRNVDYRKNDMTLFAAGNEPNGGTTISAPGTAKNAI 179

QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTIILSARSLAPDSSF 240
 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTIILSARSLAPDSSF 240

DB 180 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTIILSARSLAPDSSF 239

QY 241 WANHSKYAWMGTSMTPIVAGNVAQLREHFVNKRGITPKPSLLKALIAAGADVGLGY 300
 241 WANHSKYAWMGTSMTPIVAGNVAQLREHFVNKRGITPKPSLLKALIAAGADVGLGY 300

DB 240 WANHSKYAWMGTSMTPIVAGNVAQLREHFVNKRGITPKPSLLKALIAAGADVGLGY 299

QY 301 PNGOGGWRVTLKSLNVAAYNNESSALSTQKATYFTATAGKPKLSLWSDAPASTTA 360
 301 PNGOGGWRVTLKSLNVAAYNNESSALSTQKATYFTATAGKPKLSLWSDAPASTTA 360

DB 300 PNGOGGWRVTLKSLNVAAYNNESSALSTQKATYFTATAGKPKLSLWSDAPASTTA 359

QY 361 SYLVNDLDLVITAPNGTRYVNDFSAPFDNNMDGNNVENVFINSPOSGTYTIEVOAYN 420
 361 SYLVNDLDLVITAPNGTRYVNDFSAPFDNNMDGNNVENVFINSPOSGTYTIEVOAYN 420

DB 360 SYLVNDLDLVITAPNGTRYVNDFSAPFDNNMDGNNVENVFINSPOSGTYTIEVOAYN 419

QY 421 VPVGQPQFSLAIYN 434
 421 VPVGQPQFSLAIYN 434

DB 420 VPVGQPQFSLAIYN 433

RESULT 6
 Q9AQR2 PRELIMINARY; PRT; 433 AA.

AC Q9AQR2; 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROOC.
 OS Bacillus sp. Y.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC NCBI_TaxID=133779;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y.
 RA MEDLINE=20566675; PubMed=1118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RL Biochem Biophys Res Commun 279:313-319(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046404; BAB21267.1; -;
 DR HSSP; Q45670; 1DBI.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolyase; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 433
 SQ SEQUENCE 433 AA; 45587 MW; B81291A803C775AE CRC64;

Query Match 89.1%; Score 1999.5; DB 2; Length 433;
 Best Local Similarity 88.0%; Pred. No. 1.5e-107;
 Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGIDTGRNDSNHEAFRGKITLALYALGRTN 60
 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGIDTGRNDSNHEAFRGKITLALYALGRTN 60

DB 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGIDTGRNDSNHEAFRGKITLALYALGRTN 60

QY 61 NANDINGHGHVAGSVLNGATNKGAPQANLVFQSIMSSGGLGGLPENTLTFEQAN 120
 61 NANDINGHGHVAGSVLNGATNKGAPQANLVFQSIMSSGGLGGLPENTLTFEQAN 120

DB 61 NANDINGHGHVAGSVLNGATNKGAPQANLVFQSIMSSGGLGGLPENTLTFEQAN 119

QY 121 AGARHTNSMGAAYNGAYTTDSRNVDYRKNDMTLFAAGNEPNGGTTISAPGTAKNAI 180
 121 AGARHTNSMGAAYNGAYTTDSRNVDYRKNDMTLFAAGNEPNGGTTISAPGTAKNAI 180

DB 120 AGARHTNSMGAAYNGAYTTDSRNVDYRKNDMTLFAAGNEPNGGTTISAPGTAKNAI 179

QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTIILSARSLAPDSSF 240
 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTIILSARSLAPDSSF 240

DB 180 TVGATENLRPSFGSYADNINHVAQFSSRGPTKGRIKPDVMAPTIILSARSLAPDSSF 239

QY 241 WANHSKYAWMGTSMTPIVAGNVAQLREHFVNKRGITPKPSLLKALIAAGADVGLGY 300
 241 WANHSKYAWMGTSMTPIVAGNVAQLREHFVNKRGITPKPSLLKALIAAGADVGLGY 300

DB 240 WANHSKYAWMGTSMTPIVAGNVAQLREHFVNKRGITPKPSLLKALIAAGADVGLGY 299

QY 301 PNGOGGWRVTLKSLNVAAYNNESSALSTQKATYFTATAGKPKLSLWSDAPASTTA 360
 301 PNGOGGWRVTLKSLNVAAYNNESSALSTQKATYFTATAGKPKLSLWSDAPASTTA 360

DB 300 PNGOGGWRVTLKSLNVAAYNNESSALSTQKATYFTATAGKPKLSLWSDAPASTTA 359

QY 361 SYLVNDLDLVITAPNGTRYVNDFSAPFDNNMDGNNVENVFINSPOSGTYTIEVOAYN 420
 361 SYLVNDLDLVITAPNGTRYVNDFSAPFDNNMDGNNVENVFINSPOSGTYTIEVOAYN 420

DB 360 SYLVNDLDLVITAPNGTRYVNDFSAPFDNNMDGNNVENVFINSPOSGTYTIEVOAYN 419

QY 421 VPVGQPQFSLAIYN 434
 421 VPVGQPQFSLAIYN 434

DB 420 VPVGQPQFSLAIYN 433

RESULT 7
 Q8T9W1 PRELIMINARY; PRT; 1825 AA.

AC Q8T9W1; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Serine protease/ABC transporter TagD.
 GN TagD.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 CC NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ax4;
 RA Anjard C., Loomis W.F.;
 RT "Evolution of the ABC transporters of Dictyostelium.";
 RT submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; AF46309; AAL74253.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001140; ABC_TM_transp.

DR InterPro: IPR003439; ABC transporter.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00664; ABC_membrane_1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR ProDom: PD000006; ABC transporter; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 DR ATP-binding; Protease; Transport.
 SQ SEQUENCE 1825 AA; 202641 MW; E28160BC78613A3B CRC64;

Query Match 22.7%; Score 509.5; DB 5; Length 1825;
 Best Local Similarity 27.7%; Pred. No. 86-21;
 Matches 163; Conservative 79; Mismatches 153; Indels 193; Gaps 22;

QY 19 IYGGQGIYVADTGLDTR--NDS-----SMEAFRGKITLALYALGRINANDTNGH 68
 DB 327 LRGKQGISLADTGLDGHCFPSDSNNPIPVNSVNLNRKVTYIGSL--HDNEDYVDGH 384
 QY 69 GTHVGSVLGN-----ATKGMAPQANLVFQSI-WDSGGGLGLPSNLQTLFSQAFS 120
 DB 385 GTHVCGSAAGAPEDSLAISFSGLATDKAKIAFDLASPSNNEFVPEPDYSQLYQPLYN 444
 QY 121 AGARHITNSMGA---AVNGAYTDSRVNDYVRKN-DMTILFAAGNEGPNNGTISAPGT 175
 DB 445 AGARVHGDMSGLSIGYGSYSDAGSIDFLYHPDILIRAAAGNEQVSLSLS-QAT 503
 QY 176 AKKATVGAENTLRPF-----
 DB 504 AKKAVITVGAEGTTHSEYTTDALEYSNFEVASKSTNSLCQSPDDRYCTYTTAOCCTEXT 192
 QY 193 -----GSYAD-----NINVAQFSSRGPTKGRKIPDVMAFGTYILSA 230
 DB 564 VNGLSGCTSYIKNSYASIFSSQPELYNNENICFSSKQPTDGLKLDIVAPQYILISA 623
 QY 231 RSLA-----PDSSFWANDSKAYMGCTSMATPIYAGVAQLREH----- 271
 DB 624 RSNAGATTDOCGDGLPNTALLSE-----SGTSMATPLATATTLIKQYLVGYPT 676
 QY 272 --FVKRGIPTPKSLKALILAGADVGLGYP-----NGNMGWRV 310
 DB 677 GSIVSENKIQPTGSLIKALMINNAQLNGTFPLSNTNPSNAVDPFAGANFVQWGL 736
 QY 311 TLDSKINVAYNSS-----ALSTSQKATYTF----- 338
 DB 737 RSEEWL---YVESSGYKPRPSRWVIGELGDKKASNMWEYSLSTGQWVSYCFYKPPSS 793
 QY 339 --ATAKRP-LKISLVMSDAPASTTASVTLVNDLVIT-----APNSTRVVG-ND 384
 DB 794 GNSGSGIPIVATLVITDPSPYSGAKLVNNNDLMTNTSESEFTFYNSGSSSYNGTKG 853
 QY 385 FSAFPMNDGRRNVENVF---INSPQGTYTEVQAVNVVPGPNFS 429
 DB 854 TLPLQ---DSINVEGIITPINTKSEISFPIAGTNIPIGPONFS 898

RESULT 8

Q9GNT7 PRELIMINARY; PRT; 1702 AA.
 AC Q9GNT7
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE TAGA.
 GN TAGA.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
 CX NCBI_TaxID=44689;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Good J.R., Cabral M., Kuspa A.;
 RT "Taga, a putative serine protease/ABC transporter of Dictyostelium
 RT that is expressed at the onset of development and is required for the
 RT differentiation of a subpopulation of prespore cells.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: AF263455; AAC11416.1; .
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
 DR GO: GO:0000166; F:nucleotide binding; IEA.
 DR GO: GO:0004289; F:subtilase activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABC_TM_transp.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR ProDom: PD000006; ABC transporter; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
 DR ATP-binding; Transport.
 SQ SEQUENCE 1702 AA; 187103 MW; 4A67716303CB7131 CRC64;

Query Match 19.7%; Score 443; DB 5; Length 1702;
 Best Local Similarity 25.2%; Pred. No. 56-17;
 Matches 137; Conservative 97; Mismatches 169; Indels 140; Gaps 21;

QY 18 GLYGGQGIYVADTGLDTR--NDSMEAFRGKITLALYALGRINANDTNGHGVAG 74
 DB 300 GIKDGEIVGACDTGIDINHCFFYDINPTIGSTRKILIS-YSSGNGQIDRIDGHGTHVAG 358
 QY 75 SVLGNATN-----KGMAPQANLVFQSIWDSGGGLGLPSNLQTLFSQAFSGARIRHN 128
 DB 359 TLIGSTTVDPVSVEFSGAPNSKVAFVDLVQVSGNGLSIOSLTALYQSTYDQNAVHCD 418
 QY 129 SWGAIVGAYTDSRVNDYVRKN-DMTILFAAGNEGPNNG---TISAPGTAKNATV 182
 DB 419 AMNSNIGPFTYGTETMDRFQWHPDPLVVRSAQN-VNPFGRSYTLTGQESTSKSLVY 477
 QY 183 GATENLRPSFGSYADNIN-----
 DB 478 GSN--QPS-STYLSLDYWDMDPIVNSIRTSVCTGQSGSYGTGSDVPTQTTSDIQ 534
 QY 202 -----VAQFSSRGPTKGRKIPDVMAFGTYILSA 230
 DB 535 CCSNPILAKICCTETIQOQYQNTSTYSEFIPSLFSGVGTSGRKLPLDLAGSPILIS 594
 QY 231 RSLAPDSSFWANH-----SKYAYMGCTSMATPIYAGVAQLRE-----HVKR 276
 DB 595 R-SLGPSTI--NHCSPTISGIATSLALIAHEGSSQAAVAATSAVLYRQYRDGYFINCK 651
 QY 277 -----GITPKSLKALILAGAA---DVG.LGYENGQGWGRVTLDSKINVAYNSS--- 325
 DB 652 VNSVGFQPSASLVKALINTASINDSTLEY---SQGFNIDLSKLITTTNAQTSLDI 708
 QY 326 -----ALSTSQKATYTFATAGKPKLSLVMSDAPASTTASVTLVNDLVITTA-P 375
 DB 709 PSSIEKADPIINGETNYSYCFSLDSKADIDITLVITDPASPLSTFLVNNDLALAFV 768
 QY 376 NG--TRYVNDPSAFPENN---WDGNNVENVFINSPOGTYTEVQAVNVVPGPNFS 429
 DB 769 DGLSTISGNS-ETIFNTSQVIFDQLNVEVIRIDAPIGSIDVKIFGINIVIPNOSIS 827
 QY 430 LAI 432
 DB 828 VVI 830

RESULT 9

Q8UB0C9 PRELIMINARY; PRT; 654 AA.
 ID 08UB0C9
 AC Q8UB0C9
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Alkaline serine protease.
 GN PF1670.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Pyrococcus.
 CX NCBI_TaxId=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE010265; AAL81794.1; -
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004289; F:subtilase activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002029; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SSR; 1.
 DR KEGG: K00001; Serine protease.
 SQ SEQUENCE 654 AA; 70230 MW; 1CB145A5F505D34 CRC64;

Query Match 18.8%; Score 423; DB 17; Length 654;
 Best Local Similarity 29.6%; Pred. No. 2e-16;

Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

QY 18 GLVGGQIVAVADTGLDTRGNDSSMEAFRGKITLALYALGRINNAN-----DTNGHGH 71
 DB 154 GYDSSGTTIGIDTIGID-----ASHEDLOGKV-----IGWDFVNGRSPYDHDHGH 202
 QY 72 VAGSVLNGATN-----KMAPQANLVFQSIM--DSSGGGLGGLPSNLQTLFSAFSAARI 125
 DB 203 VASIAAGTGAASNGKYKMAPGAKLAGIKVLAGDSGSSITIKGYEMAVNDKDKIKV 262
 QY 126 HTNSWGA-----AVNGAYTTDSRNVDDYVRKQDMTILFPAAGNEGPGGTTISAP 173
 DB 263 INLSLGSQSSDGTALDQAIVNAWDA-----GLVVVAAGNSGPKYITIGSP 310
 QY 174 GTANALITVGATELRLSPFSYADNINVAOFSSRPGTKGRKPDVMAFGTYILSARS 233
 DB 311 AASAKVITLVG-----VDKIDVITFSRSGRLADGRKPEVAVAGNMTIAARS 359
 QY 234 LAPDSSFWANHDSKYAMGTSMAPIVAGNVAQLREHFYKNGKIPK--PSILKALTA 291
 DB 360 ---GTSMGQPIINDYTAAPGTSMAPIVAGIALLLQ-----AHPMTPEKXTALIE 409
 QY 292 GA-----ADVGLGVNGNGMGRTVLDKSLNAVAYVNESSLSQQA-----TYTFT 338
 DB 410 TADIVKDELADYVGA-----GRVNAVKAIV--YDNTAKLVFTGYVANKSGQTHQFV 460
 QY 339 ATAGKPKLISLVSDAPASTTASVTLVNDLVTAPNGTRGVAGNDFSAFDNMMDGRN 398
 DB 461 ISGASFYATLTYMDNAN-----SDLDLYLTPNGNQ-----VQYSTAYYG 501
 QY 399 VENVFINSPOSQTYTIEVQAVNPVGPQNFSLAIV 434
 DB 502 FEKVGYNPTDGTWTKIVSVS---GSANYQVAVS 534

RESULT 10

Q8RB02

Q8RB02 PRELIMINARY; PRT; 561 AA.

Q8RB02
 ID Q8RB02
 AC Q8RB02
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Subtilisin-like serine proteases.
 GN APR22 OR TTE0824.
 OS Thermobacteriobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermobacteriales;
 CC Thermobacteriobacteriaceae; Thermobacteriobacter.
 CX NCBI_TaxId=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4 / JCM 11007;
 RA MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Lai X., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome."
 RL Genome Res. 12:689-700(2002).
 DR EMBL: AE013049; AAM24081.1; -
 DR GO: GO:0005509; F:calcium ion binding; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004289; F:subtilase activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000209; Peptidase_S8.
 DR InterPro: IPR007280; PPC.
 DR InterPro: IPR009020; Protease_inhib.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR Pfam: PF04151; PPC; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SSR; 1.
 DR KEGG: K00001; Serine protease.
 SQ SEQUENCE 561 AA; 59968 MW; BA9CC52F7083A18 CRC64;

Query Match 18.6%; Score 418; DB 16; Length 561;
 Best Local Similarity 30.9%; Pred. No. 3.2e-16;

Matches 140; Conservative 66; Mismatches 157; Indels 90; Gaps 19;

QY 6 GIVRADVAQSSYGYGGQIVAVADTGLDTRGNDSSMEAFRGKITLALYALGRINNAN-- 63
 DB 155 GILTK---ARSDFGYTGKNTITAIIDTIGIDGNHYLS-----GGKI-----IGKCDPLNKK 201
 QY 64 ---DTNGHGHVAVSVLNGATN-----KMAPQANLVFQSIMDS--SGGLGLGPSNLQTL 114
 DB 202 TTPYDDNHGHVAVSVLNGATN-----KMAPQANLVFQSIMDS--SGGLGLGPSNLQTL 114
 QY 115 FSQAFSAGARHTNSWGAIVNGAYTTDSRNVDDYVRKQDMTILFPAAGNEGPGGTTISAP 174
 DB 262 VQNDVYGIKVINLSLSTGTSSTGDTSTSLAVNRVADSGIVVVAAGNSGPARYTIGSP 321
 QY 175 TAKNAITVGATENLRSPFSYADNINVAOFSSRPGTKGRKIPDVMAFGTYILSARS 234
 DB 322 AAEKAIIVAMADV-----GELGRFL---ASFSSRGPADRIKEDIAAPRINNTAK--- 371
 QY 235 APDSSFWANHDSKYAMGTSMAPIVAGNVAQLREHFYKNGKIPKPSILKALTAGAA 294
 DB 372 ---ANSYNGSVYVTSSTGTSMAPIVAGTVALMLN---ANPILTPDA---KIIINSTAK 419
 QY 295 DVLGVNGNGMGRTVLDKSLNAVAYVNESSLSQQA-----TYTFT 339
 DB 420 SMGPPSKNVDYAGRLGLEYAIRAGNFRGNIDVPHNYIT--SGILPGRYSDTTFNA 477
 QY 340 T-AGKPKLISLV--WSDAPASTTASVTLVNDLVTAPNGTRGVAGNDFSAFDNMMDG 395
 DB 478 TMTSYPIAITLILIDMANYNP-----DFDIYLYDPSGTLIKS-----TG 517
 QY 396 RNVENVFINSPOSQTYTIEVQAVNPVGPQNF 428

Db 518 TORSETITLIPSOITGYVVKVYSYR---GSGNY 547

RESULT 11

Q9FEB24 PRELIMINARY; PRT; 1239 AA.
 ID Q9FEB24
 AC Q9FEB24
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative secreted peptidase.
 GN SC07188 OR SC8A11.16C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
 NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Redebach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,
 RA Knaeht H., Hopwood D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=97000351; Pubmed=8843436;
 RA Redebach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,
 RA Knaeht H., Hopwood D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=2196410; Pubmed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 RN EMBL; AL939130; CAC01588.1; -
 DR HSSP; Q98405; IMPT.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001317; PA.
 DR InterPro; IPR002029; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Complete proteome.
 SQ SEQUENCE 1239 AA; 128505 MW; 8F5E9AC6E8B1260A CRC64;

Query Match 18.2%; Score 407.5; DB 16; Length 1239;
 Best Local Similarity 33.3%; Pred. No. 3.7e-15;
 Matches 133; Conservative 40; Mismatches 145; Indels 73; Gaps 12;
 QY 8 VKADVAOSSY-----GLYGGQIVAVADTGLDTGRNDSMHEAFRGKITLALYAG 57

Db 219 VEADLADSTAOIGAPRPMAGNTGQVAVLDTGVAG-----HPDLADRIARQSFV 272

QY 58 RTNNANDTNGHTVAGSVLGNATN---KMAPQANLVFQSIMDSGGGLPSNLQT 113
 Db 273 PDENTDDRDGHGTVAAGTIGTGAASAGKEGVAPGARLIGKVLDNS -GRGQISWTIAA 331
 QY 114 LFSQAFSAGARIHNNMSGAAVNGAVTTD -SRNVDDYVRKQDMTLFLPAGNEGPGNTISA 112
 Db 332 MEMAIVEERHAKIVNNSIGSGEQSDSPMSRAVDRLSAQGFVVAAGN -GGEAGSITGA 390
 QY 173 PGTAKNAITVGATENTLPSFGSVADNINHYAOFSSRSGPTDGRGIRKPDVAPGTYLISARS 232
 Db 391 PGVATSAALTVA-----VDATDLAPSSQCFRVDGALKPEITAPGVGIIAA-- 437
 QY 233 SLADDSFWANHDSKAVVMGTSMAATIVAGNVAQLREHVNRKRGITPKRSLLKAL--- 289
 Db 438 ---NSSPAGNGAVYSLGTSNATPHVAGAAALL-----AARPDLSGSLKQV 484
 QY 290 IAGADVGLGYPNQNGMGRVTLDPKSLN-----VAYVNESSALSTQKATYFTPATAG 342
 Db 485 LASSHRTPRDARQAGSRVDVAARAGVVASATAPAGSSPGPYRBLVYTTNTTGA 544
 QY 343 KPLKISLVMSDA-----PASTTASVTLVND 367
 Db 545 VTLELSVAATHPAGVFRLSASRYTVPAGHTADVTLTID 583

RESULT 12

Q9ENV1 PRELIMINARY; PRT; 430 AA.
 ID Q9ENV1
 AC Q9ENV1
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Intracellular alkaline serine proteinase.
 GN O82375
 OS Oceanobacillus thelyensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
 RX MEDLINE=2220767; Pubmed=12235376;
 RA Takami H., Takaki Y., Uchiyama T.;
 RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments.";
 RL Nucleic Acids Res. 30:3927-3935(2002).
 RN EMBL; AP004601; BAC14311.1; -
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002029; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Complete proteome.
 SQ SEQUENCE 430 AA; 45838 MW; 6D09A99BCEB310F CRC64;

Query Match 17.2%; Score 387; DB 16; Length 430;
 Best Local Similarity 35.1%; Pred. No. 1.4e-14;
 Matches 118; Conservative 53; Mismatches 125; Indels 40; Gaps 16;

QY 2 DVARGIVKADVAOSSYGLYGGQIVAVADTGLDTGRNDSMHEAFRKIT--ALYALGRT 59
 Db 121 DTASSINADVAKES -GLTGQGSTIAVIDTGIHP-----HEDLEGRITGFADPVKQT 172
 QY 60 NNADTNGHTVAGSVLGNATN---KMAPQANLVFQSIMDSGGGLPSNLQTLF 115
 Db 173 EPYDD -NGHGTGAGDAAGNALSDDQYGGPAPDANLVGVKLVNKGTS -GSLSTVIEGI- 229
 QY 116 SQAFSAGARITTNWGAANGAAYTTDSRNVDDYVR-----KNDMTLPPAAGNEGPGGT 169

```

Db      230 DWICNOSKYNINILISTL-GSDATEPAEGPVPVNAVETAMDGMVYCAAGNSGSDGT 288
Qy      170 ISAGTKMAITVATENLRPSFGSYADNINHVAGFSRPTDGRKPKVMAFGYIIS 229
      289 VGSFGISPKVITVAAADNNNTAERS---DDVAEFSRSRPTIDGLTKPPLTTPGVDIYS 344
Qy      230 ARS--SLAPSSFWANHDSKYAVMGTSMAPIVAGNVAOLREHFVNRGITPKPSILKA 287
      345 LRAPGSFIDKTKNSARVGSVNYISLSTGSMATPICAGIVAGLQ---SDSLT--PNQVKE 399
Qy      288 ALIAGADVGLGYNGNGMGRTLDKSLNVAAYNE 323
      400 KLMACODLGQS-PN-VQAGAYL--NAANLININE 430

RESULT 13
Q9FC06 PRELIMINARY; PRT; 1253 AA.
AC      09FC06;
DT      01-MAR-2001 (Tremblrel. 16, Created)
DT      01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT      01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE      Putative secreted peptidase.
GN      SC07176 OR SC8A11.04C.
OS      Streptomyces coelicolor.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OX      Streptomycinae; Streptomycetaceae; Streptomyces.
QX      NCBI_TaxID=1902;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      Saunders D.C., Harris D.;
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      Cardeno A.M., Parkhill J., Barrrell B.G., Rajandream M.A.;
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      MEDLINE=97000351; PubMed=8843436;
RX      Kadenbach K., Kleser H.M., Denapate D., Eichner A., Cullum J.,
      Khasni H., Hopwood D.A.;
RT      "A set of ordered cosmids and a detailed genetic and physical map for
      the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL      Mol. Microbiol. 21:77-96(1996).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RX      MEDLINE=2196410; PubMed=12000953;
RA      Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
      Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
      Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
      Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
      Huang C.H., Kleser H., Jarke L., Murthy U., Oliver K., O'Neill S.,
      Rabinowitz S., Rajandream M.A., Rutherford K., Rutter S.,
      Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
      Warren T., Wietzorrek A., Woodward J., Barrrell B.G., Parkhill J.,
      Hopwood D.A.;
RT      "Complete genome sequence of the model actinomycete Streptomyces
      coelicolor A3(2).";
RL      Nature 417:141-147(2002).
RN      EMBL; A0939130; CAC01576.1; -.
DR      HSSP; Q99405; IMPT.
DR      GO; GO:0008233; F:peptidase activity; IEA.
DR      GO; GO:0004289; F:subtilase activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      InterPro; IPR001317; PA.
DR      InterPro; IPR000209; Peptidase_S8.
DR      Pfam; PF02225; PA; 1.
DR      Pfam; PF00082; Peptidase_S8; 1.

```

```

DR      PRINTS; PR00723; SUBTILISIN.
DR      PROSITE; PS00136; SUBTILASE_ASP; 1.
DR      PROSITE; PS00137; SUBTILASE_HIS; 1.
DR      PROSITE; PS00138; SUBTILASE_SER; 1.
DR      Complete proteome.
SQ      SEQUENCE 1253 AA; 130971 MW; AA69B417EEEDB89 CRC64;

Query Match      17.1%; Score 385; DB 16; Length 1253;
Best Local Similarity 29.4%; Pred. No. 7,4e-14;
Matches 140; Conservative 52; Mismatches 162; Indels 122; Gaps 19;

Qy      16 SYGLYGQGIIVAAVDTGLDTRNDSSMHEAFRGKITALLALGRTNANDNGHGTVAAGS 75
      233 SGGVTGEGVAVAVLDGTVDAG-----HPDFAGRIATAFVFPDQVTDNRNGHTVAAT 266
      76 VLNGKATN---KMAPOANLVFQSIIDSSG-----GLGG 106
      287 VAGTGAASGVKEKVAPEASLHIGKVLIDNSGSGDSQSVTLAEMENAVRDQAKIVSNLGD 346
      107 LPSNLOTLPQAFSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGN 166
      347 SPTDGTDLSEAV-----NWLASFTGA-----LFVAAAGNGEPE 380
Qy      167 GGTISAPGTAKNATTVGATENTLRPSFGSYADNINHVAGFSRGP-TRDGRIKPDVAAFGT 225
      381 AYVGTTPAADAALTVAGVNG--FGKG---VDQLDFSSRGPRVGDNAVKDPLTAPGV 433
Qy      226 YILARSLSLAPSSFWANHDSKYAVMGTSMAPIVAGNVAOLREHFVKRKGITPKPSIL 285
      434 GVLAARSRVAREG-----EGAYQSLGTSMAPIVAGNVAALAAEHPDWTG-----QRL 482
Qy      286 KAALIAADVGLGYNGNGMGRTLDKSLNVAAYNESSALSTSQ-----KATYT 336
      483 KEALVGTAGTQRSP-FDAGSGRVDAAVRSTLNSGDAFCAHAPYTPGQTVRRDVT 541
Qy      337 FTATAGKPKLSTVWSDA-----PASTASVTLVNDLDTVTPANG--- 377
      542 YTNAGPAVALDIALSTARELPEGLFTLSEAVYVPAIGTASVGVITILD--AAEDNGAYA 599
Qy      378 TRYV--GNDS---APFDNNMGRRNVENVFI---NSPQSGTYIT-EVQAVNVP 422
      600 TLVVASGADGAVLARTEPVGVNKEGRATLTALTAKDHHKPLSGVILKQVERNTAP 655
Db

RESULT 14
P95684 PRELIMINARY; PRT; 1102 AA.
AC      P95684;
DT      01-MAY-1997 (Tremblrel. 03, Created)
DT      01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT      01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE      Subtilisin-like protease.
OS      Streptomyces albobacillus.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OX      Streptomycinae; Streptomycetaceae; Streptomyces.
QX      NCBI_TaxID=1887;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S-1253;
RX      MEDLINE=97144528; PubMed=8990295;
RA      Suzuki M., Teguchi S., Yamada S., Kojima S., Miura K., Momose H.;
RT      "A novel member of the subtilisin-like protease family from
      Streptomyces albobacillus.";
RL      J. Bacteriol. 179:430-438(1997).
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
RN      EMBL; D83672; BAAL2040.1; -.
DR      HSSP; P00782; 28BT.
DR      MEROPS; S08.069; -.
DR      GO; GO:0008233; F:peptidase activity; IEA.
DR      GO; GO:0004289; F:subtilase activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      InterPro; IPR002860; GH_BNR.
DR      InterPro; IPR000209; Peptidase_S8.

```

DR Pfam; PF00212; BNR. 2.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Hydrolyase; Protease; Serine protease.
 SQ SEQUENCE 1102 AA; 114128 MW; F9E4AD2590FE559E CRC64;

Query Match 16.3%; Score 365; DB 2; Length 1102;
 Best Local Similarity 31.3%; Pred. No. 8.8e-13;
 Matches 146; Conservative 56; Mismatches 177; Indels 88; Gaps 19;

QY VAR-----GIKAVVAGS-----SYLVGGQIVAVADTGLDGNDSMEAFRG 48
 DB 183 VAAVWLDGVKASLDTSVQIGTPKMEAGYDCKGVKIALDLDGVD-----ATHEDLG 236
 QY 49 KITALVALGRNNANDTNGHGHVAGSVLGNCA---TNKGMAPQANTLVFQSIMDSGGL 104
 DB 237 QVTASKNFTSAPTGGDVVGHGTHVASLAAGTGAQSGKTGYVPAKILNGKVLDDAG-- 294
 QY 105 GGLPSNLQITFSQAFSAGARIFHNSGAAVNGAYTTDSRNVDYRK--NDMTILF--AA 160
 DB 295 FGDDSGILACMEWAAAGADIYVMSLG---GMDTPEITPLEAAVDKLSAEKGIIFALAA 350
 QY 161 GNEGPNGTISAPGTAKNATTVGATENLRPSFGSYADNINHYAQPSSRGP-TRDGRKPD 219
 DB 351 GNGSPQ--SIGSPGASDALTVA-----VDDCKLADPSSTPRIGDGVKPD 397
 QY 220 VMAPGYILSARSSLPADSSFWANHDSKYAVNGTSMATPIVAGVAOLREHFVNRGIT 279
 DB 398 LTPAGVDITPAASAKGNDIAKEVEKEPAGWTTISGTMATPHVAGAALLKQHP----- 452
 QY 280 PKSSLKAALIAAGADVGLGYPNGNQGMRVTLDSLVNAYVNESALS----- 328
 DB 453 WKRAELKGLALTSTKYG--KITPFEQSGRQVQDKAITQTVAEPVSLFVQVQPHADD 510
 QY 329 --TSQKATYTFATAGKPKISLVMSD-----APAS--TTASVTVVNDLVLITAP--NGT 378
 DB 511 KPYTKLTYRNLGTEDVTLKLTSTATGPKKAPAFGLGASTL-----TYPANGT 562
 QY 379 RYVGNDFSAFPDNNMGRNNVENVFINSPS-----GTITIEQANV 421
 DB 563 ASVDVTAIDTRLGAGVDGYSAVVATGAGQSVRTAAVAREVESYV 609

RESULT 15
 Q82B14 PRELIMINARY; PRT; 1208 AA.
 AC Q82B14;
 DT 01-UN-2003 (TEMBLrel. 24, Created)
 DT 01-UN-2003 (TEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Putative peptidase.
 GN SAV5721.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.,
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:528-531(2003).
 DR EMBL; AP005044; BAC73433.1;
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001337; PA.
 DR InterPro; IPR002009; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Complete proteome.
 SQ SEQUENCE 1208 AA; 125548 MW; E650B5E3ABE512B CRC64;

Query Match 16.1%; Score 360.5; DB 16; Length 1208;
 Best Local Similarity 31.5%; Pred. No. 1.8e-12;
 Matches 129; Conservative 53; Mismatches 161; Indels 67; Gaps 17;

QY 8 VKAVVAGS-----YGLVGGQIVAVADTGLDGNDSMEAFRGKITALVALG 57
 DB 186 VEAQMASSNQIGTAAADAGLTDGVTAVVLDIGVDT-----THPDLAGRVSRSKFT 239
 QY 58 RTNNANDTNGHGHVAGSVLGNCA---TNKGMAPQANTLVFQSIMDSGGLGGLPSNLQIT 113
 DB 240 DGEVADRNGHGHVHTVTSVGSGLASDCTERGVAPGATTLAVGVLSQGA---GSESQT 235
 QY 114 LFSQAFSA---GARIFHNSGGA--AVNGAYTTD--SRNVDYRKNDMTILFLAAGNEPN 166
 DB 296 IAGMEMARVRARVATVMSLSGSTEASDG---TTPMAEAVDTLSEETGALFVVAAGNTGA- 351
 QY 167 GGTISAPGTAKNATTVGATENLRPSFGSYADNINHYAQPSSRGPTR-DGRKPDVWAPGT 225
 DB 352 PSSISGPGADSLTVA-----VDSDDRAAYFTSAGPRGDNALKEDLAAFGV 400
 QY 226 YILSARSSLPADSSFWANHDSKYAVNGTSMATPIVAGVAOLREHFVNRGITPRESL 285
 DB 401 DIRAARQLAPGTGY-----YTSMSGTSMATPHVAGVAAALAEQHPDWTGARLKDALM 453
 QY 286 KAA--LTAGADVGLGYPNGNQGNG--RVTLDSLVNAYVNESALSTSQKATYTFATAG 342
 DB 454 STSEQLDASVYQLAGRVSVPDAVGARVATGSGADLGFHRWPHDADRPVTKVTYISNSD 513
 QY 343 KPLKISLVMSDAPASTASVTLVNDLVLITAP---NGTRYVGNDFSAF 388
 DB 514 TVVELSLAVGAPAGV---ATLA---DTALTVAHGTAAITVVGDSKAP 557

Search completed: March 31, 2004, 16:08:56
 Job time : 35.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.1304 Seconds

(without alignment)
2595.843 Million cell updates/sec

Title: US-09-985-689a-3

Sequence: 1 NDVARGIVKADVAQNNYGLY.....EYQAVNPSPGPRSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A_Geneseq_29Jan04.*

1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003s.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2263	100.0	433	5	AAW50082
2	2251	99.5	433	5	AAW50084
3	2239	98.9	433	2	AAW26274
4	2239	98.9	433	2	AAW61495
5	2239	98.9	433	2	AAW95698
6	2239	98.9	433	3	AAW9207
7	2239	98.9	433	3	AAW44619
8	2234	98.7	433	5	AAW50083
9	2230	98.5	433	5	AAW89548
10	2040	90.1	433	5	AAW50086
11	2040	90.1	433	5	AAW89547
12	2020.5	89.3	434	2	AAW50085
13	2006.5	88.7	434	5	AAW50081
14	1998.5	88.3	639	2	AAW17089
15	1995.5	88.2	640	2	AAW17091
16	1994.5	88.1	434	5	AAW50080
17	1994.5	88.1	434	2	AAW17090
18	1948.5	86.1	639	2	AAW17087
19	1948.5	86.1	640	2	AAW17088
20	1837.5	81.2	434	5	AAW50090
21	1514	66.9	345	2	AAW62230
22	1514	66.9	345	2	AAW21654
23	443	19.6	659	2	AAW24121
24	443	19.6	659	2	AAW94840
25	422.5	18.7	412	2	AAW94836

26	422.5	18.7	522	2	AAW24122	AAW24122	Pyrococcus
27	422.5	18.7	522	2	AAW94838	AAW94838	Hyperther
28	422.5	18.7	654	2	AAW24129	AAW24129	Pyrococcus
29	422.5	18.7	654	2	AAW94841	AAW94841	Hyperther
30	401.5	17.7	659	2	AAW24123	AAW24123	Protease
31	370.5	16.4	545	4	ABW99483	ABW99483	Protease
32	366.5	15.3	1237	6	ABU11343	ABU11343	Protein e
33	345.5	15.3	1079	4	ABW1180	ABW1180	Transglut
34	345.5	15.3	1079	6	ABU07391	ABU07391	Foreign p
35	343.5	15.2	520	2	AAW13666	AAW13666	Fragment
36	343.5	15.2	734	2	AAW13667	AAW13667	Streptomy
37	343.5	15.2	823	2	AAW13668	AAW13668	DHPA-mel
38	312	13.8	806	2	AAW27481	AAW27481	RP-III re
39	305	13.5	903	2	AAW87007	AAW87007	Hyperther
40	305	13.5	1398	2	AAW87008	AAW87008	Pyrococce
41	305	13.5	1398	2	AAW24124	AAW24124	Pyrococce
42	305	13.5	1398	2	AAW94839	AAW94839	MO9856926
43	296.5	13.1	580	7	ADD24927	ADD24927	Xanthomon
44	295	13.0	519	6	ABP76735	ABP76735	Streptomy
45	295	13.0	19938	6	ABP76678	ABP76678	Streptomy

ALIGNMENTS

RESULT 1
ID AAW50082 standard; protein; 433 AA.

AAW50082;
AC AAW50082;

DT 12-AUG-2002 (first entry)

XX Bacillus sp D6-(FERM P1592) alkaline protease protein fragment.

XX Alkaline protease, detergent, laundry, bleaching, dishwasher.

XX Bacillus sp.

XX EP1209233-A2.

XX 29-MAY-2002.

XX 22-NOV-2001; 2001EP-00127851.

XX 22-NOV-2000; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.

XX (KAO) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

XX PI Okuda M, Saeki K;

XX WPI; 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions.

XX Claim 5; Page 13-15; 25pp; English.

XX This invention describes novel Bacillus sp. alkaline proteases useful in

XX detergent compositions, especially in laundry, bleaching or automatic

XX dishwasher detergents. The novel proteases have an increased detergent %

XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

XX sequence represents a fragment of the alkaline protease E-1 from Bacillus

XX sp strain D6-(FERM-P1592) described in the method of the invention

XX Sequence 433 AA;

XX Query Match 100.0%; Score 2263; DB 5; Length 433;

XX Best Local Similarity 100.0%; Pred. No 2.4e-170; Indels 0; Gaps 0;

XX Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 NDVARGIVKADVAQNNYGLYGGQVAVADTGLDTGRNDSMHAFFKRTALVALGRIN 60

```

Db      1 NDVARGIVKADVQNNNGLYGGQVVAADTGLDTRGNDSSMHEARFGKITLALYLGRTN 60
Qy      61 NADPNHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLTLFSGQAMNA 120
Db      61 NADPNHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLTLFSGQAMNA 120
Qy      121 GARHTNSWGAAPVNGATYANSRQVDEVYRNNMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121 GARHTNSWGAAPVNGATYANSRQVDEVYRNNMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Qy      181 VGATENYRPFSGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPSSFW 240
Db      181 VGATENYRPFSGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPSSFW 240
Qy      241 ANYSKYAYVGGTSMATPIVAGNVAQLREHFIKRGITTPPSLIKALTAGATDVGIGYP 300
Db      241 ANYSKYAYVGGTSMATPIVAGNVAQLREHFIKRGITTPPSLIKALTAGATDVGIGYP 300
Qy      301 SGDOGGRVTLDDKSLNVAAYNEATALTGOKATYSFOTQAGKPKISLWTDAPGSTTAS 360
Db      301 SGDOGGRVTLDDKSLNVAAYNEATALTGOKATYSFOTQAGKPKISLWTDAPGSTTAS 360
Qy      361 YTLVNDLDLVITLTPNGOKTYGNDFSYPDNMDGRNNVENVFNAPOSGTYTIEVOAYNV 420
Db      361 YTLVNDLDLVITLTPNGOKTYGNDFSYPDNMDGRNNVENVFNAPOSGTYTIEVOAYNV 420
Qy      421 PSGPQRFSLAIYH 433
Db      421 PSGPQRFSLAIYH 433

```

RESULT 2
AAMS0084
ID AAMS0084 standard; protein; 433 AA.

```

XX      12-AUG-2002 (first entry)
XX      DE Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.
XX      KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX      OS Bacillus sp.
XX      PN EP1209233-A2.
XX      PD 29-MAY-2002.
XX      PF 22-NOV-2001; 2001EP-00127851.
XX      PR 22-NOV-2000; 2000JP-00355166.
XX      PR 12-APR-2001; 2001JP-00114048.
XX      PA (KAO) KAO CORP.
XX      PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX      PI Okuda M, Saeki K;
XX      DR WPI; 2002-437518/47.
XX      PT New modified alkaline proteases useful in detergent compositions.
XX      PS Claim 5; Page 16-18; 25pp; English.
XX      CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX      CC detergent compositions, especially in laundry, bleaching or automatic
XX      CC dishwasher detergents. The novel proteases have an increased detergency &
XX      CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX      CC sequence represents a fragment of the alkaline protease SD-521 from
XX      CC Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the
XX      CC invention

```

```

XX      SQ Sequence 433 AA;
Qy      Query Match 99.5%; Score 2251; DB 5; Length 433;
Qy      Best Local Similarity 99.3%; Pred. No. 2,2e-169;
Qy      Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVQNNNGLYGGQVVAADTGLDTRGNDSSMHEARFGKITLALYLGRTN 60
Db      1 NDVARGIVKADVQNNNGLYGGQVVAADTGLDTRGNDSSMHEARFGKITLALYLGRTN 60
Qy      61 NADPNHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLTLFSGQAMNA 120
Db      61 NADPNHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLTLFSGQAMNA 120
Qy      121 GARHTNSWGAAPVNGATYANSRQVDEVYRNNMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121 GARHTNSWGAAPVNGATYANSRQVDEVYRNNMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Qy      181 VGATENYRPFSGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPSSFW 240
Db      181 VGATENYRPFSGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPSSFW 240
Qy      241 ANYSKYAYVGGTSMATPIVAGNVAQLREHFIKRGITTPPSLIKALTAGATDVGIGYP 300
Db      241 ANYSKYAYVGGTSMATPIVAGNVAQLREHFIKRGITTPPSLIKALTAGATDVGIGYP 300
Qy      301 SGDOGGRVTLDDKSLNVAAYNEATALTGOKATYSFOTQAGKPKISLWTDAPGSTTAS 360
Db      301 SGDOGGRVTLDDKSLNVAAYNEATALTGOKATYSFOTQAGKPKISLWTDAPGSTTAS 360
Qy      361 YTLVNDLDLVITLTPNGOKTYGNDFSYPDNMDGRNNVENVFNAPOSGTYTIEVOAYNV 420
Db      361 YTLVNDLDLVITLTPNGOKTYGNDFSYPDNMDGRNNVENVFNAPOSGTYTIEVOAYNV 420
Qy      421 PSGPQRFSLAIYH 433
Db      421 PSGPQRFSLAIYH 433

```

RESULT 3
AAR26274
ID AAR26274 standard; protein; 433 AA.

```

XX      AAR26274;
XX      DT 05-FEB-1993 (first entry)
XX      DE Alkali-protease Ya enzyme.
XX      OS Alkali resistance; surface active agent resistance; detergency improver.
XX      PN JP04197182-A.
XX      PD 16-JUL-1992.
XX      PF 28-NOV-1990; 90JP-00327110.
XX      PR 28-NOV-1990; 90JP-00327110.
XX      PA (LION) LION CORP.
XX      PI WPI; 1992-288440/35.
XX      DR N-PSDB; AAQ27516.
XX      PT DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
XX      PT resistance and improves detergency.
XX      PS Claim 2; Page 1; 17pp; Japanese.
XX      CC The sequence is that of alkali-protease Ya enzyme which can be used in

```

CC the recombinant production of Ya enzyme. Ya enzyme is excellent in alkali
 CC resistance and surface active agent resistance and improves detergency
 XX
 SQ Sequence 433 AA;

Query Match 98.9%; Score 2239; DB 2; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;
 Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIIMDSGGIGLPSNLNTLFSQANNA 120
 DB 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIIMDSGGIGLPSNLNTLFSQANNA 120
 QY 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNIHAFSSRGATRDGRITKPDVTAGTTLASGSIAPDSFW 240
 DB 181 VGATENYRPSFGSIADNPNIHAFSSRGATRDGRITKPDVTAGTTLASGSIAPDSFW 240
 QY 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPKPSLTKALIAAGTDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPKPSLTKALIAAGTDVGLGYP 300
 QY 301 SGDQGWGRVTLDSLNVAAYNENATALTGQKATYSFOAQGPILKISLWTDAPGSTTAS 360
 DB 301 SGDQGWGRVTLDSLNVAAYNENATALTGQKATYSFOAQGPILKISLWTDAPGSTTAS 360
 QY 361 YTLVNDLDELIVITAPNGQKYVGNDFSYPDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 DB 361 YTLVNDLDELIVITAPNGQKYVGNDFSYPDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 4
 AAM61495 standard; protein; 433 AA.
 XX
 AC AAM61495;
 XX
 DT 06-NOV-1998 (first entry)
 XX
 DE Modified Bacillus lichen protease.
 XX
 KM Bacillus lichen Y protease; polyethylene glycol; PEG; soap;
 KM methoxypolyethylene glycol; MPEG; skin; hair care product; cosmetic;
 KM lipstick; hair gel; sun oil; shampoo; hair dye; insect repellent.
 OS Bacillus sp.
 XX
 FH Key
 FT Modified-site 1.433
 FT /note= "The enzyme is modified by methoxypolyethylene glycol
 FT molecules covalently attached to the N-terminal amino
 FT group and to fourteen unspecified amino groups of lysine
 FT residues present on the surface of the enzyme"
 XX
 EN MO9830682-A1.
 XX
 PD 16-UTL-1998.
 XX
 PF 12-JAN-1998; 98WO-DK000015.
 XX
 PR 10-JAN-1997; 97DK-00000038.
 PR 25-JUN-1997; 97DK-00000754.

XX (NOVO) NOVO-NORDISK AS.
 PA
 XX
 PI Olsen AA, Prento A;
 XX
 DR WPI; 1998-399132/34.
 XX
 PT New enzyme modified by attachment of many polymeric molecules - useful in
 PT skin and hair care products, has reduced tendency to cause sensitisation
 PT and increased stability.
 PS Claim 16; Page 44-45; 56pp; English.
 XX

CC The present sequence represents the Bacillus lichen Y protease. The
 CC invention claims for enzymes covalently modified on their surface by the
 CC attachment of a large number of small polymeric molecules, e.g.
 CC polyethylene glycols (PEG). The polymeric molecules are coupled to the N-
 CC terminal amino group and the amino groups of lysine residues found on the
 CC surface of the enzyme. In the example given, the lichen Y protease was
 CC modified using methoxypolyethylene glycol (MPEG) as the polymeric molecule.
 CC The N-terminal amino group and the amino groups of the fourteen lysine
 CC residues present on the surface of the lichen Y protease were modified.
 CC Modification of the enzymes increases the stability and/or reduces the
 CC sensitising potential (allergenicity) of the enzyme, without
 CC significantly reducing enzymatic activity. Also, using a large number of
 CC relatively small polymeric molecules, rather than a few very large ones,
 CC provides a more even effect with reduced activity loss. The modified
 CC enzymes are claimed to be useful as components of a wide range of skin
 CC and hair care products, e.g. soaps, cosmetics, creams, gels, lipsticks,
 CC hair gels, sun oils, shampoos, hair dyes, insect repellants, etc
 XX

Query Match 98.9%; Score 2239; DB 2; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;
 Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIIMDSGGIGLPSNLNTLFSQANNA 120
 DB 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIIMDSGGIGLPSNLNTLFSQANNA 120
 QY 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNIHAFSSRGATRDGRITKPDVTAGTTLASGSIAPDSFW 240
 DB 181 VGATENYRPSFGSIADNPNIHAFSSRGATRDGRITKPDVTAGTTLASGSIAPDSFW 240
 QY 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPKPSLTKALIAAGTDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPKPSLTKALIAAGTDVGLGYP 300
 QY 301 SGDQGWGRVTLDSLNVAAYNENATALTGQKATYSFOAQGPILKISLWTDAPGSTTAS 360
 DB 301 SGDQGWGRVTLDSLNVAAYNENATALTGQKATYSFOAQGPILKISLWTDAPGSTTAS 360
 QY 361 YTLVNDLDELIVITAPNGQKYVGNDFSYPDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 DB 361 YTLVNDLDELIVITAPNGQKYVGNDFSYPDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 5
 AAM95698 standard; protein; 433 AA.
 ID AAM95698

XX AAK95698;
 AC 16-JUN-1999 (first entry)
 DT Bacillus sp. Lion Y protease.
 DE Bacillus sp. Lion Y protease.
 XX PD498; subtilisin; Lion Y; protease; skin-care; feed; additive; soap;
 KW cosmetic; hair dye; sunscreen; acne; antiperspirants; insect repellent;
 KW deodorant; detergent; food; breadmaking; textile-treating; oral; dermal;
 KW pharmaceutical; agrochemical.
 XX Bacillus sp.
 OS WO9900489-A1.
 PN 07-JAN-1999.
 PD 22-JUN-1998; 98WO-DK000270.
 XX 25-JUN-1997; 97DK-00000753.
 PR 07-JUL-1997; 97US-0051830P.
 XX (NOVO) NOVO-NORDISK AS.
 PA Olsen AA, Fatum TM, Deussen H, Roggen EL;
 PI WPI; 1999-095735/08.
 DR New modified polypeptide with attached low molecular weight polymer - has
 PT reduced respiratory allergenicity; useful in skin care products.
 PT detergents, as food additives or textile-treating compositions.
 XX Claim 10; Page 48-49; 60pp; English.
 PS The sequence is that of Lion Y protease. This can be used as an active
 XX ingredient: (i) in personal care products (especially skin-care products
 CC such as soaps, cosmetics, hair dyes, sunscreens, anti-lacne products,
 CC antiperspirants, insect repellants or deodorants); (ii) in detergents (as
 CC laundry, dishwashing or hard-surface cleaners); (iii) food or feed
 CC additives (e.g. for breadmaking); (iv) in textile-treating compositions,
 CC or (v) in oral or dermal pharmaceuticals and agrochemicals
 XX
 XX Sequence 433 AA;
 SQ

Query Match 98.9%; Score 2239; DB 2; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;
 Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 60
 DB 1 NDVARGIVADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 60
 QY 61 NANDPBGHGTVAAGSVLGNALNKGMAPQANLVFQSIMSSGGIGLPNLTLPFGQANNA 120
 DB 61 NANDPBGHGTVAAGSVLGNALNKGMAPQANLVFQSIMSSGGIGLPNLTLPFGQANNA 120
 QY 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNGSPNSGTISAGTAKNALT 180
 DB 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNGSPNSGTISAGTAKNALT 180
 QY 181 VGATENTYRRSPFGSIADNPVHIAQFSSRGATRDGRTRKPDVTAGPTITLARSGLAPDSFW 240
 DB 181 VGATENTYRRSPFGSIADNPVHIAQFSSRGATRDGRTRKPDVTAGPTITLARSGLAPDSFW 240
 QY 241 AATNSKYAAMGGSMTPIVAGNVAQLREHFIKNGITPSPSLIKALILAGATDVGLGYP 300
 DB 241 AATNSKYAAMGGSMTPIVAGNVAQLREHFIKNGITPSPSLIKALILAGATDVGLGYP 300
 QY 301 SEDQMGKRVTLDKSLNVAAYNEATLTTGQAKTYSFQTAQKPLKISLWTDAPGSTTAS 360
 DB 301 SEDQMGKRVTLDKSLNVAAYNEATLTTGQAKTYSFQTAQKPLKISLWTDAPGSTTAS 360
 QY 360 GSDQMGKRVTLDKSLNVAAYNEATLTTGQAKTYSFQTAQKPLKISLWTDAPGSTTAS 360
 DB 360 GSDQMGKRVTLDKSLNVAAYNEATLTTGQAKTYSFQTAQKPLKISLWTDAPGSTTAS 360

QY 361 YTLVNDLDELVTAPNGOKTVANDFSYPDNNDGRRNVEFINAPQSGTITIEVOAYNV 420
 DB 361 YTLVNDLDELVTAPNGOKTVANDFSYPDNNDGRRNVEFINAPQSGTITIEVOAYNV 420
 QY 421 PSGPQRFSLAIYH 433
 DB 421 PSGPQRFSLAIYH 433
 RESULT 6
 AAY69207
 ID AAY69207 standard; protein; 433 AA.
 XX AAY69207;
 AC 30-MAY-2000 (first entry)
 DT Amino acid sequence of protease Lion Y.
 XX Protease Lion Y; polypeptide-polymer conjugate; washing performance;
 KW respiratory allergenicity; allergic reaction; detergent formulation;
 KW laundry; dishwashing; hard surface cleaner; agricultural chemical;
 KW skin care; cosmetic; oral pharmaceutical; dental pharmaceutical;
 KW textile processing.
 XX Bacillus sp.
 OS WO200004138-A1.
 PN 27-JAN-2000.
 PD 16-JUL-1999; 99WO-DK000406.
 XX 17-JUL-1998; 98DK-00000951.
 PR (NOVO) NOVO-NORDISK AS.
 PA Bauditz P, Fatum TM, Olsen AA, Deussen H, Petersen DA;
 PI WPI; 2000-195024/17.
 DR New polypeptide-polymer conjugate, particularly enzyme conjugate, useful
 PT in detergent formulations.
 PT Disclosure; Page 74-76; 79pp; English.
 PS The present sequence represents a protease Lion Y protein. The protein
 CC may be used to produce the conjugates of the invention. The specification
 CC describes polypeptide-polymer conjugates which have improved washing
 CC performance and reduced respiratory allergenicity compared with the
 CC unconjugated polypeptide. The polymer provides, in water, a conformation
 CC that shields the molecular surface effectively, preventing association of
 CC antibodies that can induce an allergic reaction. The conjugates are used
 CC in industrial compositions, particularly detergent formulations (laundry,
 CC dishwashing or hard surface cleaners), but also in agricultural
 CC chemicals, skin care products (cosmetics and toiletries), oral and dental
 CC pharmaceuticals, or textile processing and treatment compositions
 XX
 XX Sequence 433 AA;
 SQ

Query Match 98.9%; Score 2239; DB 3; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;
 Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 60
 DB 1 NDVARGIVADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 60
 QY 61 NANDPBGHGTVAAGSVLGNALNKGMAPQANLVFQSIMSSGGIGLPNLTLPFGQANNA 120
 DB 61 NANDPBGHGTVAAGSVLGNALNKGMAPQANLVFQSIMSSGGIGLPNLTLPFGQANNA 120
 QY 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNGSPNSGTISAGTAKNALT 180
 DB 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNGSPNSGTISAGTAKNALT 180

Db 121 GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFPAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLPDSSFW 240
 Db 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLPDSSFW 240
 QY 241 ANYNSKYAVMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLTKAALIAAGTVDGLGYP 300
 Db 241 ANYNSKYAVMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLTKAALIAAGTVDGLGYP 300
 QY 301 SDDQGWGRVTLDKSLNVAAYVNEATLTTGOKATYSFQOAGKPKLSLWTDAPGSTTAS 360
 Db 301 SDDQGWGRVTLDKSLNVAAYVNEATLTTGOKATYSFQOAGKPKLSLWTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGOKYVGNDFSYPDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGOKYVGNDFSYPDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIVH 433
 Db 421 PSGPQRFSLAIVH 433
 RESULT 7
 AAY44619
 ID AAY44619 standard; protein; 433 AA.
 AC AAY44619;
 XX 07-APR-2000 (first entry)
 DT 07-APR-2000 (first entry)
 DE Bacillus lion y enzyme.
 XX
 XX Lion Y enzyme; protease; allergic response; industrial composition;
 KM co-polymer; ethylene oxide; EO; propylene oxide; PO; conjugate;
 KM allergenicity; detergent; cosmetic; toiletries; textile treatment;
 KM agrochemical; pharmaceutical; food; feed additive.
 XX
 OS Bacillus sp.
 XX
 PN WO9967370-A1.
 PD 29-DEC-1999.
 XX
 PF 23-JUN-1999; 99WO-DK000359.
 PR 23-JUN-1998; 98DK-00000809.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 PA
 PI Deussen H, Olsen AA, Fatum TM, Roggen EL;
 XX
 XX WPI; 2000-136981/12.
 DR
 PT New conjugate of polypeptide, especially an enzyme, with copolymer of
 PT ethylene oxide and propylene oxide, used in e.g. cleaning compositions,
 PT has reduced allergenicity.
 XX
 PS Claim 10; Page 56-57; 62pp; English.
 XX
 XX The present sequence is a Bacillus lion y enzyme, which is a protease
 CC capable of inducing an allergic response upon inhalation. The enzyme can
 CC be covalently coupled to a co-polymer comprising ethylene oxide (EO) and
 CC propylene oxide (PO) to reduce its allergenicity. This enzyme-polymer
 CC conjugate can be used in industrial compositions such as detergents,
 CC cosmetics, toiletries, textile treatment compositions, agrochemicals,
 CC oral and dermal pharmaceuticals and food and feed additives
 XX
 XX Sequence 433 AA;
 Query Match 98.9%; Score 2239; DB 3; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;

Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNNGYLGQGOVAVVADTGLDTGRNDSWHEAFRGKITLVALGRTN 60
 Db 1 NDVARGIVKADVAQNNNGYLGQGOVAVVADTGLDTGRNDSWHEAFRGKITLVALGRTN 60
 QY 61 NADPNHGHTHVAAGSVLGNALNKGMAPOANLVFQSIMDSGGLGGLPSNLNTLFGQAMVA 120
 Db 61 NADPNHGHTHVAAGSVLGNALNKGMAPOANLVFQSIMDSGGLGGLPSNLNTLFGQAMVA 120
 QY 121 GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFPAAGNEGPNSGTISAPGTAKNAIT 180
 Db 121 GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFPAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLPDSSFW 240
 Db 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLPDSSFW 240
 QY 241 ANYNSKYAVMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLTKAALIAAGTVDGLGYP 300
 Db 241 ANYNSKYAVMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLTKAALIAAGTVDGLGYP 300
 QY 301 SDDQGWGRVTLDKSLNVAAYVNEATLTTGOKATYSFQOAGKPKLSLWTDAPGSTTAS 360
 Db 301 SDDQGWGRVTLDKSLNVAAYVNEATLTTGOKATYSFQOAGKPKLSLWTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGOKYVGNDFSYPDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGOKYVGNDFSYPDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIVH 433
 Db 421 PSGPQRFSLAIVH 433
 RESULT 8
 AAM50083
 ID AAM50083 standard; protein; 433 AA.
 AC AAM50083;
 XX 12-AUG-2002 (first entry)
 DT 12-AUG-2002 (first entry)
 DE Bacillus sp Y-(FERM BP-1029) alkaline protease protein fragment.
 XX
 KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS Bacillus sp.
 XX
 PN EP1209233-A2.
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-00127851.
 PR 22-NOV-2000; 2000UP-00355166.
 PR 12-APR-2001; 2001JP-00114048.
 XX
 XX (KAO) KAO CORP.
 PA
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX
 XX WPI; 2002-437518/47.
 DR
 PT New modified alkaline proteases useful in detergent compositions.
 PT
 PS Claim 5; Page 15-16; 25pp; English.
 XX
 XX This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergent
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

CC sequence represents a fragment of the alkaline protease Ya from *Bacillus*
CC sp strain Y-(FERM BP-1029) described in the method of the invention
XX
SQ Sequence 433 AA;

Query Match 98.7%; Score 2234; DB 5; Length 433;
Best Local Similarity 98.6%; Pred. No. 4.8e-168;
Matches 427; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDTRNDSSMHEAFRGKITLVALGRIN 60
DB 1 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDTRNDSSMHEAFRGKITLVALGRIN 60
QY 61 NANDPENGHTHVASVTLGNALNKGMAPQANLVFQSIMDSSGGLGSPNLNTLFGQANNA 120
DB 61 NANDPENGHTHVASVTLGNALNKGMAPQANLVFQSIMDSSGGLGSPNLNTLFGQANNA 120
QY 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFILSARSSLAPOSSFW 240
DB 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFILSARSSLAPOSSFW 240
QY 241 ANYNSKYAVMGSTSMATPIVAGNVAQLREHFIKNRGITPKPSLIKALILAGATDVLGYP 300
DB 241 ANYNSKYAVMGSTSMATPIVAGNVAQLREHFIKNRGITPKPSLIKALILAGATDVLGYP 300
QY 301 SGGQGWGRVTLDSLVNAVYNNENTALTGGKATYSFGTQAGKPLKISLWTDAPGSTTAS 360
DB 301 SGGQGWGRVTLDSLVNAVYNNENTALTGGKATYSFGTQAGKPLKISLWTDAPGSTTAS 360
QY 361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
DB 361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
QY 421 PSGQRFSLAIIVH 433
DB 421 PSGQRFSLAIIVH 433

RESULT 9

AAW89548
ID AAW89548 standard; protein; 636 AA.

XX AC AAW89548;
XX 12-APR-1999 (first entry)
XX DE *Bacillus* sp. alkaline protease Y.
XX KM Alkaline protease Y; detergent; surfactant; leather processing;
XX OS *Bacillus* sp.
XX PN WO9856927-A2.
XX PD 17-DEC-1998.
XX PF 09-JUN-1998; 98WO-US012005.
XX PR 12-JUN-1997; 97US-00873479.
XX PA (NOVO) NOVO NORDISK BIOTECH INC.
XX PI Sioma A, Christianson L;
XX DR WPI, 1999-080908/07.
XX PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
PT dishwashing detergents and for leather processing.

XX
XX Claim 3; Page 55-56; 77pp; English.

CC This is the amino acid sequence of a *Bacillus* sp. alkaline protease Y
CC that is said to have good alkali and surfactant resistance and improved
CC detergency. It shows 77% identity to a newly isolated protease (see
CC AAW89547) of *Bacillus* sp. Jp170 (NCIB 12513). The invention provides
CC vectors, recombinant host cells and methods for the recombinant
CC production of such proteases. The protease are used in laundry and
CC dishwashing detergents, for institutional and industrial cleaning, and
CC for leather processing, as well as for debittering and enhancing the
CC degree of hydrolysis of protein hydrolysates, for flavour development
CC through hydrolysis of peptides, degradation of undesired peptides and in
CC enzymatic synthesis of peptides. They have enhanced stability towards
CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
CC peroxyl type. The invention also provides mutant cells in which the
CC protease activity is diminished. Such cells can be used for the
CC production of heterologous recombinant proteins

XX
SQ Sequence 636 AA;

Query Match 98.5%; Score 2230; DB 2; Length 636;
Best Local Similarity 98.6%; Pred. No. 1.7e-167;
Matches 427; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDTRNDSSMHEAFRGKITLVALGRIN 60
DB 204 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDTRNDSSMHEAFRGKITLVALGRIN 263
QY 61 NANDPENGHTHVASVTLGNALNKGMAPQANLVFQSIMDSSGGLGSPNLNTLFGQANNA 120
DB 61 NANDPENGHTHVASVTLGNALNKGMAPQANLVFQSIMDSSGGLGSPNLNTLFGQANNA 120
QY 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFILSARSSLAPOSSFW 240
DB 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFILSARSSLAPOSSFW 240
QY 241 ANYNSKYAVMGSTSMATPIVAGNVAQLREHFIKNRGITPKPSLIKALILAGATDVLGYP 300
DB 241 ANYNSKYAVMGSTSMATPIVAGNVAQLREHFIKNRGITPKPSLIKALILAGATDVLGYP 300
QY 301 SGGQGWGRVTLDSLVNAVYNNENTALTGGKATYSFGTQAGKPLKISLWTDAPGSTTAS 360
DB 301 SGGQGWGRVTLDSLVNAVYNNENTALTGGKATYSFGTQAGKPLKISLWTDAPGSTTAS 360
QY 361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
DB 361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
QY 421 PSGQRFSLAIIVH 433
DB 421 PSGQRFSLAIIVH 433

RESULT 10

AAW50086
ID AAW50086 standard; protein; 433 AA.

XX AC AAW50086;
XX 12-AUG-2002 (first entry)
XX DE *Bacillus* sp alkaline protease protein A-2 fragment.
XX KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX OS *Bacillus* sp.
XX XX EPI209233-A2.

Query Match	90.1%; Score 2040; DB 5; Length 433;
Best Local Similarity	88.9%; Pred. No. 1e-152;
Matches	385; Conservative 24; Mismatches 24; Indels 0; Gaps 0
QY	1 NDVARGIVKADVQNNYGYGGQGVVAADGDLTGGRNDSMHEAFRGKITALYALGRIN 60
DB	1 NDVARGIVKADVQNNYGYGGQGVVAADGDLTGGRNDSMHEAFRGKITALYALGRIN 60
QY	61 NADPNRGHGTIVAGSVLGNALNKKMAPQANLVFGSINDSGGLGCLPSNLNTFSQAMNA 120
DB	61 NADPNRGHGTIVAGSVLGNALNKKMAPQANLVFGSINDSGGLGCLPSNLNTFSQAYSA 120
QY	121 GARHTNSWGAFAVNGAYTANSROVDEYVRANDMTVLFPAAGNEGSGSTISAPGTAKAAT 180
DB	121 GARHTNSWGAFAVNGAYTANSROVDEYVRANDMTVLFPAAGNEGSGSTISAPGTAKAAT 180
QY	181 VGATENYRPFSGSLANPNHIAOFSSRGATDGRFKPVTNAPGFTILASRSLAPSSFTW 240
DB	181 VGATENYRPFSGSLANPNHIAOFSSRGATDGRFKPVTNAPGFTILASRSLAPSSFTW 240
QY	241 ANYNSKAYAVGSGTSMATPIVAGNVQOLREHPFKNGKITPKPSLTKAALIGATDVLGYF 300
DB	241 ANYNSKAYAVGSGTSMATPIVAGNVQOLREHPFKNGKITPKPSLTKAALIGATDVLGYF 300
QY	301 SGGQGGKRTVLDKSLNVAAYVNEATLTTGQKATYSFOQDAKPKLILVNTDAPSGTTS 360
DB	301 SGGQGGKRTVLDKSLNVAAYVNEATLTTGQKATYSFOQDAKPKLILVNTDAPSGTTS 360
QY	361 YTVLNDLDELVTAPNGQKTVGNDFSFYVDNNMDGNNENVFINAPQSGTYTIEVQAYNV 420
DB	361 YTVLNDLDELVTAPNGKTKVGNDFAPYDNNMDGNNENVFINAPQSGTYTIEVQAYNV 420
QY	421 PSQGPQFSLAIIVH 433
DB	421 PSQGPQFSLAIIVH 433

DE	Bacillus jpl170 protease.
XX	
KM	Protease; detergent; surfactant; leather processing; debittering;
XX	flavour.
OS	
XX	Bacillus sp.
FH	
FT	Key Location/Qualifiers
FT	Peptide 1..33 /note= "signal peptide"
FT	Region 34..208 /note= "preprio region"
FT	Protein 209..641 /note= "mature protein"
PN	
XX	MO9856927-A2.
PD	
XX	17-DEC-1998.
PF	
XX	09-JUN-1998; 98WC-USO12005.
PR	
XX	12-JUN-1997; 97US-00873479.
PA	(NOVO) NOVO NORDISK BIOTECH INC.
PI	
XX	Sloma A, Christianson L;
DR	
XX	WPI; 1999-080908/07.
DR	N-PSDB; AA82382.
XX	
PT	Novel protease from Bacillus subtilis LC20 - useful in laundry and
PT	dishwashing detergents and for leather processing.
PB	Claim 7; Page 53-54; 77bp; English.
XX	
CC	This is the amino acid sequence of a novel protease of Bacillus sp. JPl170
CC	(NCBI 12513), as deduced from the nucleotide sequence of an isolated gene
CC	(see AA82382). The entire protein, including the signal peptide and
CC	preprio region, has 77% identity to alkaline protease Y (see AA89548)
CC	from Bacillus. The invention provides vectors, recombinant host cells and
CC	methods for the recombinant production of the protease. The protease is
CC	used in laundry and dishwashing detergents, for institutional and
CC	industrial cleaning, and for leather processing, as well as for
CC	debittering and enhancing the degree of hydrolysis of proteins,
CC	hydrolysates, for flavour development through hydrolysis of peptides.
CC	Degradation of undesired peptides and in enzymatic synthesis of peptides.
CC	It has enhanced stability towards oxidation under alkaline conditions,
CC	e.g. towards bleaching agents of the peroxy type. The invention also
CC	provides mutant cells in which the protease activity is diminished. Such
CC	cells can be used for the production of heterologous recombinant proteins
XX	
SQ	Sequence 641 AA:
Query Match	90.1%; Score 2040; DB 2; Length 641;
Beet Local Similarity	88.9%; Pred. No.1.8e-15;
Matches 385; Conservative	24; Mismatches 24; Indels 0; Gaps 0;
QY	1 NDVARGIYKADVAQNNGVLGYQGQGVAVAVDTGLDTRGDSMEHEAFRGKITALVALGRTN 60
Db	209 NDVARGIYKADVAQNNGFLGYQGQGVAVAVDTGLDTRGRDSMEHEAFRGKITALVALGRTN 268
QY	61 NANPBNHGCTTVAGSVTGNAANKMAPOANLVFQSIMSSGGTGLPENLMTLTFEQAWNNA 120
Db	269 NANPBNHGCTTVAGSVTGNAANKMAPOANLVFQSIMSSGGTGLPENLMTLTFEQAWNSA 328
QY	121 GARHTNSWGAPVNGAVATANSROYDEVYRNNDMITVLPAAQNEGPNSGTISAPGTAKNAIT 180
Db	329 GARHTNSWGAPVNGAVATTDSRRNDVDYVRKKDMITLFFAAGNEGPGSGTISAPGTAKNAIT 388
QY	161 VGATNTVPSPSGSTADNPNTHAQSSRPATGDGRIPKVTPAPGFITLSABASLPADSFPW 240
Db	389 VGATNTVPSPSGSTADNPNTNHVAQSSRKPTDGRIKPVPVAPGVTILARSGLPADSSEFW 448

QY 241 ANNSKAYVGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAAGATDVLGY 300
DB 449 ANHDSKAYVGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAAGATDVLGY 508
QY 301 SGDOGWRVTLDKSLNVAAYNEATALTGOKATYSPFOTQAGKPLKISLWMTAPGSTTA 360
DB 509 NGNGQWRVTLDKSLNVAAYNEATALTGOKATYSPFOTQAGKPLKISLWMTAPGSTTA 568
QY 361 YTLVNDLDLVITAPNGOKTVGNDPSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYN 420
DB 569 LTLVNDLDLVITAPNGOKTVGNDPSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYN 628
QY 421 PPSGQRFSLATVH 433
DB 629 PPSGQRFSLATVH 641

RESULT 12
ID AAM50085 standard; protein; 434 AA.
XX AAM50085;
XX AC
XX 12-AUG-2002 (first entry)
XX DE Bacillus sp alkaline protease protein A-1 fragment.
XX KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX OS Bacillus sp.
XX PN EPI209233-A2.
XX PD 29-MAY-2002.
XX PF 22-NOV-2001; 2001EP-00127851.
XX PR 22-NOV-2000; 2000JP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX PA (KAOS) KAO CORP.
XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX PI Okuda M, Saeki K;
XX DR WPI; 2002-437518/47.
XX PS New modified alkaline proteases useful in detergent compositions.
XX PS Claim 5; Page 18-19; 25pp; English.
XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency %
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease A-1 from Bacillus
XX CC sp NCIB1289 described in the method of the invention
XX SQ Sequence 434 AA;

Query Match 89.3%; Score 2020.5; DB 5; Length 434;
Best Local Similarity 88.2%; Pred. No. 3.7e-151;
Matches 383; Conservative 28; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLGGQGVAVADTGLDTGRNDSMHEARRGKITATYALGRN 60
DB 1 NDVARGIVKADVAQNNYGLGGQGVAVADTGLDTGRNDSMHEARRGKITATYALGRN 60
QY 61 NADPNHGTHVAGSVLGNAL-NKGAAPQANLVFQSIIMSSGGLGIPSNLNTLFSQAWN 119
DB 61 NADPNHGTHVAGSVLGNAL-NKGAAPQANLVFQSIIMSSGGLGIPSNLNTLFSQAWN 120
QY 120 AGARIHNSMGAPVNGAYTANSRQVDEYVANNMTYVFAAGNGRPSGTISAPGTAKNAI 179

DB 121 AGARIHNSMGAPVNGAYTANSRQVDEYVANNMTYVFAAGNGRPSGTISAPGTAKNAI 180
QY 180 TVGATENRPSFGSIAONPNHIAQFSSRGATPDGRIPKDTAPQFTLSRSSIAPSSSF 239
DB 181 TVGATENRPSFGSIAONPNHIAQFSSRGATPDGRIPKDTAPQFTLSRSSIAPSSSF 240
QY 240 WANNYSKAYVGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAAGATDVLGY 299
DB 241 WANNYSKAYVGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAAGATDVLGY 300
QY 300 PPSGQWRVTLDKSLNVAAYNEATALTGOKATYSPFOTQAGKPLKISLWMTAPGSTTA 359
DB 301 PPSGQWRVTLDKSLNVAAYNEATALTGOKATYSPFOTQAGKPLKISLWMTAPGSTTA 360
QY 360 SYTLVNDLDLVITAPNGOKTVGNDPSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYN 419
DB 361 SYTLVNDLDLVITAPNGOKTVGNDPSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYN 420
QY 420 VPSGQRFSLATVH 433
DB 421 VPSGQRFSLATVH 434

RESULT 13
ID AAM50081 standard; protein; 434 AA.
XX AAM50081;
XX AC
XX 12-AUG-2002 (first entry)
XX DE Bacillus sp KSM-KP9860 alkaline protease protein fragment.
XX KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX OS Bacillus sp.
XX PN EPI209233-A2.
XX PD 29-MAY-2002.
XX PF 22-NOV-2001; 2001EP-00127851.
XX PR 22-NOV-2000; 2000JP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX PA (KAOS) KAO CORP.
XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX PI Okuda M, Saeki K;
XX DR WPI; 2002-437518/47.
XX PS New modified alkaline proteases useful in detergent compositions.
XX PS Claim 5; Page 12-13; 25pp; English.
XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency %
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease KP9860 from
XX CC Bacillus sp strain KSM-KP9860 described in the method of the invention
XX SQ Sequence 434 AA;

Query Match 88.7%; Score 2006.5; DB 5; Length 434;
Best Local Similarity 88.2%; Pred. No. 4.7e-150;
Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLGGQGVAVADTGLDTGRNDSMHEARRGKITATYALGRN 60

XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX WPI, 1999-287736/27.
DR N-PSDB, AAX37279.
XX

PT Alkali protease from *Bacillus* used in washing powders.

PS Disclosure, Page 63-68; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of
CC *Bacillus*. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX

SQ Sequence 640 AA;

Query Match 88.2%; Score 1995.5; DB 2; Length 640;

Best Local Similarity 87.6%; Pred. No. 6e-149;
Matches 380; Conservative 26; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIYKADVAQNNYGLVGGQVAVADTGLDTGRNDSMHEAFKGTITAYALGRIN 60
DB 207 NDVARGIYKADVAQSSYGLVGGQIVAVADTGLDTGRNDSMHEAFKGTITAYALGRIN 266
QY 61 NADPNHGHTVAGSVIGN-ALNKMAPQANLVFQSIMDSGGIGLPSNYLTLFQAWN 119
DB 267 NADPNHGHTVAGSVIGNSTNKMAPQANLVFQSIMDSGGIGLPSNYLTLFQAWN 326
QY 120 AGARHTNSWGAIPNGAYTANSROYDEYVANNMTVTFAAGTEGPNSTISAPGTAKNAI 179
DB 327 AGARHTNSWGAIPNGAYTANSROYDEYVANNMTVTFAAGTEGPNSTISAPGTAKNAI 386
QY 180 TVGATENYRPSFGSIADNPNIHIAQFSRSGATRDGRIKPDVTAPGTFTLSARSSLAPDSF 239
DB 387 TVGATENYRPSFGSIADNPNIHIAQFSRSGATRDGRIKPDVTAPGTFTLSARSSLAPDSF 446
QY 240 WANYNSKAYMGTSMTPIVAGNVAQLREHFIKNGITPKPSLIKALITAGATDVGLGY 299
DB 447 WANYNSKAYMGTSMTPIVAGNVAQLREHFIKNGITPKPSLIKALITAGATDVGLGY 506
QY 300 PSGDQGWGRVTLDSLVNAVYNEATALTGQXATVSFQTQAGKPLKISLVMTDAPGSTTA 359
DB 507 PSGDQGWGRVTLDSLVNAVYNEATALTGQXATVSFQTQAGKPLKISLVMTDAPGSTTA 566
QY 360 SYTLVNDLDLVITAPNGQKTVGDNFSYPYNNMDGRNNVENVFINAPQSGTYTIEVOAYN 419
DB 567 SYTLVNDLDLVITAPNGQKTVGDNFSYPYNNMDGRNNVENVFINAPQSGTYTIEVOAYN 626
QY 420 VPSGQQRFSLAIVH 433
DB 627 VPSGQQRFSLAIVH 640

Search completed: March 31, 2004, 16:04:30
Job time : 48.1304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:59:39 ; Search time 13.9892 Seconds

(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689a-3

Perfect score: 2263
Sequence: 1 NDVARGIVKADVAQNNYGLY.....EYQAVNPSPGPRFSLATVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Parents AA:*

1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	98.9	433	4	US-09-104-623A-4
2	2239	98.9	433	4	US-09-104-623A-4
3	2239	98.9	433	4	US-09-104-623A-4
4	2239	98.9	433	4	US-09-104-623A-4
5	2239	98.9	433	4	US-09-104-623A-4
6	2239	98.9	433	4	US-09-104-623A-4
7	2239	98.9	433	4	US-09-104-623A-4
8	2239	98.9	433	4	US-09-104-623A-4
9	2239	98.9	433	4	US-09-104-623A-4
10	2239	98.9	433	4	US-09-104-623A-4
11	2239	98.9	433	4	US-09-104-623A-4
12	2239	98.9	433	4	US-09-104-623A-4
13	2239	98.9	433	4	US-09-104-623A-4
14	2239	98.9	433	4	US-09-104-623A-4
15	2239	98.9	433	4	US-09-104-623A-4
16	2239	98.9	433	4	US-09-104-623A-4
17	2239	98.9	433	4	US-09-104-623A-4
18	2239	98.9	433	4	US-09-104-623A-4
19	2239	98.9	433	4	US-09-104-623A-4
20	2239	98.9	433	4	US-09-104-623A-4
21	2239	98.9	433	4	US-09-104-623A-4
22	2239	98.9	433	4	US-09-104-623A-4
23	2239	98.9	433	4	US-09-104-623A-4
24	2239	98.9	433	4	US-09-104-623A-4
25	2239	98.9	433	4	US-09-104-623A-4
26	2239	98.9	433	4	US-09-104-623A-4
27	2239	98.9	433	4	US-09-104-623A-4

28	305	13.5	903	1	US-08-750-532-1	Sequence 1, Appl1
29	305	13.5	1398	1	US-08-750-532-9	Sequence 8, Appl1
30	305	13.5	1398	4	US-08-894-818B-8	Sequence 6, Appl1
31	305	13.5	1398	4	US-09-445-472-6	Sequence 18, Appl1
32	280.5	12.4	237	1	US-08-750-532-18	Sequence 1, Appl1
33	272	12.0	275	2	US-08-750-406A-1	Sequence 1, Appl1
34	272	12.0	275	3	US-09-327-118-1	Sequence 5, Appl1
35	269	11.9	269	1	US-08-431-387-5	Sequence 6, Appl1
36	269	11.9	269	1	US-08-322-677A-10	Sequence 10, Appl1
37	269	11.9	269	1	US-08-322-676-10	Sequence 10, Appl1
38	269	11.9	269	1	US-08-140-083A-10	Sequence 10, Appl1
39	269	11.9	269	3	US-08-898-218-10	Sequence 10, Appl1
40	269	11.9	269	3	US-08-848-793-10	Sequence 5, Appl1
41	269	11.9	269	3	US-09-255-502-5	Sequence 3, Appl1
42	269	11.9	269	3	US-09-024-532-3	Sequence 4, Appl1
43	269	11.9	269	3	US-08-269-050-4	Sequence 1, Appl1
44	269	11.9	269	3	US-08-090-207-1	Sequence 1, Appl1
45	269	11.9	269	3	US-08-090-207-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-104-623A-4
Sequence 4, Application US/09104623A

Patent No. 6303752

GENERAL INFORMATION:

APPLICANT: Olsen, Arne Agerlin

APPLICANT: Deussen, Heinz-Josef

APPLICANT: Roggen, Erwin Ludo

TITLE OF INVENTION: A Modified Polypeptide

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESS: No. 63037520 No. 63037520sk of No. 6303752th America, Inc.

STREET: 405 Lexington Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/104,623A

FILING DATE: 25-JUN-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rozek, Carol

REGISTRATION NUMBER: 36,993

REFERENCE/DOCKET NUMBER: 5256, 200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 amino acids

TYPE: amino acid

STRANDEDNESS: linear

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

STRAIN: Bacillus sp. Y

US-09-104-623A-4

Query Match

Best Local Similarity 98.8%, Pred. No. 8.9e-155;

Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGTGRNDSMEHAFKITALYALRTN 60

```

Db      1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEARFKITLALYALGRIN 60
Qy      61 NADDPNGHGTTHVAGSVLGNALNKGMAPQANLVFQSIIMSSGGIGLPSNLNTLFSQANNA 120
Db      61 NADDPNGHGTTHVAGSVLGNALNKGMAPQANLVFQSIIMSSGGIGLPSNLNTLFSQANNA 120
Qy      121 GARHHTNSWGAAPNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAGTANNAIT 180
Db      121 GARHHTNSWGAAPNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAGTANNAIT 180
Qy      181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
Db      181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
Qy      241 ANYSKXAYWGTSMATPIVAGNVAOLREHFINKRGITTPKPSLIKALLAGATDVGLGYP 300
Db      241 ANYSKXAYWGTSMATPIVAGNVAOLREHFINKRGITTPKPSLIKALLAGATDVGLGYP 300
Qy      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQKPKLSLWTDAPGSTTAS 360
Db      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQKPKLSLWTDAPGSTTAS 360
Qy      361 YTLVNDLDLVITAPNGQKYGNDPSYPDNNDGRNNVENVFINAQSGTYTIEVOAYNV 420
Db      361 YTLVNDLDLVITAPNGQKYGNDPSYPDNNDGRNNVENVFINAQSGTYTIEVOAYNV 420
Qy      421 PSGQRFSLAIHV 433
Db      421 PSGQRFSLAIHV 433

```

```

RESULT 2
US-09-019-532-4
; Sequence 4, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Olesen, Arne Agerlin
; APPLICANT: Prent, Annette
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922.204-US
; CURRENT APPLICATION NUMBER: US/09/019,532B
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0038/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-019-532-4

```

```

Query Match      98.9%; Score 2239; DB 4; Length 433;
Best Local Similarity 98.8%; Pred. No. 8.9e-165;
Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEARFKITLALYALGRIN 60
Db      1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEARFKITLALYALGRIN 60
Qy      61 NADDPNGHGTTHVAGSVLGNALNKGMAPQANLVFQSIIMSSGGIGLPSNLNTLFSQANNA 120
Db      61 NADDPNGHGTTHVAGSVLGNALNKGMAPQANLVFQSIIMSSGGIGLPSNLNTLFSQANNA 120
Qy      121 GARHHTNSWGAAPNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAGTANNAIT 180
Db      121 GARHHTNSWGAAPNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAGTANNAIT 180

```

```

Db      121 GARHHTNSWGAAPNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAGTANNAIT 180
Qy      181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
Db      181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
Qy      241 ANYSKXAYWGTSMATPIVAGNVAOLREHFINKRGITTPKPSLIKALLAGATDVGLGYP 300
Db      241 ANYSKXAYWGTSMATPIVAGNVAOLREHFINKRGITTPKPSLIKALLAGATDVGLGYP 300
Qy      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQKPKLSLWTDAPGSTTAS 360
Db      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQKPKLSLWTDAPGSTTAS 360
Qy      361 YTLVNDLDLVITAPNGQKYGNDPSYPDNNDGRNNVENVFINAQSGTYTIEVOAYNV 420
Db      361 YTLVNDLDLVITAPNGQKYGNDPSYPDNNDGRNNVENVFINAQSGTYTIEVOAYNV 420
Qy      421 PSGQRFSLAIHV 433
Db      421 PSGQRFSLAIHV 433

```

```

RESULT 3
US-09-338-746-4
; Sequence 4, Application US/09338746
; Patent No. 6638526
; GENERAL INFORMATION:
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Olesen, Arne A.
; APPLICANT: Fatum, Line M.
; APPLICANT: Røgeen, Erwin L.
; TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
; FILE REFERENCE: 5619.200-US
; CURRENT APPLICATION NUMBER: US/09/338,746
; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: PA 1998 00809
; EARLIER FILING DATE: 1998-07-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-338-746-4

```

```

Query Match      98.9%; Score 2239; DB 4; Length 433;
Best Local Similarity 98.8%; Pred. No. 8.9e-165;
Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEARFKITLALYALGRIN 60
Db      1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRGNDSSMHEARFKITLALYALGRIN 60
Qy      61 NADDPNGHGTTHVAGSVLGNALNKGMAPQANLVFQSIIMSSGGIGLPSNLNTLFSQANNA 120
Db      61 NADDPNGHGTTHVAGSVLGNALNKGMAPQANLVFQSIIMSSGGIGLPSNLNTLFSQANNA 120
Qy      121 GARHHTNSWGAAPNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAGTANNAIT 180
Db      121 GARHHTNSWGAAPNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAGTANNAIT 180
Qy      181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
Db      181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
Qy      241 ANYSKXAYWGTSMATPIVAGNVAOLREHFINKRGITTPKPSLIKALLAGATDVGLGYP 300
Db      241 ANYSKXAYWGTSMATPIVAGNVAOLREHFINKRGITTPKPSLIKALLAGATDVGLGYP 300
Qy      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQKPKLSLWTDAPGSTTAS 360
Db      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQKPKLSLWTDAPGSTTAS 360

```

Db 301 SGGQGWGRVTLDSKLNVAAYVNEATLALATGOKATYSQAQGRPLKSLVWTDAPGSTTAS 360
QY 361 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVQAINV 420
Db 361 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVQAINV 420
QY 421 PSQPGRFSLAIHV 433
Db 421 PSQPGRFSLAIHV 433

RESULT 4
US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-873-479-43

Query Match 98.8%; Score 2239; DB 2; Length 635;
Best Local Similarity 98.8%; Pred. No. 1.5e-164;
Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 241 ANVSKYAYNGGSMATPIVAGNVACUREHFICRGTTPKPSLIKALLIAGATDVGIGP 300
Db 443 ANVSKYAYNGGSMATPIVAGNVACUREHFICRGTTPKPSLIKALLIAGATDVGIGP 302
QY 301 SGGQGWGRVTLDSKLNVAAYVNEATLALATGOKATYSFQQAQGRPLKSLVWTDAPGSTTAS 360
Db 503 SGGQGWGRVTLDSKLNVAAYVNEATLALATGOKATYSFQQAQGRPLKSLVWTDAPGSTTAS 362
QY 361 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVQAINV 420
Db 563 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVQAINV 422
QY 421 PSQPGRFSLAIHV 433
Db 623 PSQPGRFSLAIHV 635

RESULT 5
US-08-873-479-42
; Sequence 42, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-873-479-42

Query Match 90.1%; Score 2040; DB 2; Length 641;
Best Local Similarity 88.9%; Pred. No. 3.4e-149;
Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;

QY 121 GARIHNSWGA PYNAGYATANSROVDEYVRNNDMTVLFAAGNPGNSGTISAPGTAKNAIT 180
DB 329 GARIHNSWGA PYNAGYATANSROVDEYVRNNDMTVLFAAGNPGNSGTISAPGTAKNAIT 388
QY 181 VGATENYRPSFGSIADNPNIHIAQFSRSGATRDGRIPDVTAPGTILSARSSLAADSSSF 240
DB 389 VGATEMLRPSFGSIADNPNIHIAQFSRSGATRDGRIPDVTAPGTILSARSSLAADSSSF 448
QY 241 ANNSKXAYWGGTSMATPIVAGNVAOLREHFINKNGITPKPSILKALTAGADVGLGP 300
DB 449 ANNSKXAYWGGTSMATPIVAGNVAOLREHFINKNGITPKPSILKALTAGADVGLGP 508
QY 301 SGDQGWGRVTLDSKSLNVAAYNEATALTGOKATYSPOTQAGKPKLSLWTDAPGSTTAS 360
DB 509 NGNQGWRVTLDSKSLNVAAYNEATALTGOKATYSPOTQAGKPKLSLWTDAPGSTTAS 568
QY 361 YTLVNDLVLITPAPNOQKYYGNDSEYFYDNNMDGRNNVENVTINAPQSTYITIEVOAYN 420
DB 569 YTLVNDLVLITPAPNOQKYYGNDSEYFYDNNMDGRNNVENVTINAPQSTYITIEVOAYN 628
QY 421 PPSGQRFSLAIYH 433
DB 629 PPSGQRFSLAIYH 641

RESULT 6

US-09-509-814A-4
; Sequence 4, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-4

Query Match 88.3%; Score 1998.5; DB 4; Length 639;
Best Local Similarity 88.0%; Pred. No. 5.3e-146;
Matches 382; Conservative 26; Mismatches 25; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVQNNYGLYGQGOVVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 60
DB 206 NDVARGIVKADVQNNYGLYGQGOVVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 265
QY 61 NADDPNGHGHVAGSVLGN-ALNKGAPOANLVFQSIIMDSGGGLGGLPSNLNTLFSQAWN 119
DB 266 NADDPNGHGHVAGSVLGN-ALNKGAPOANLVFQSIIMDSGGGLGGLPSNLNTLFSQAWN 325
QY 120 AGARIHNSWGA PYNAGYATANSROVDEYVRNNDMTVLFAAGNPGNSGTISAPGTAKNAI 179
DB 326 AGARIHNSWGA PYNAGYATANSROVDEYVRNNDMTVLFAAGNPGNSGTISAPGTAKNAI 385
QY 180 TVGATENYRPSFGSIADNPNIHIAQFSRSGATRDGRIPDVTAPGTILSARSSLAADSSSF 239

DB 386 TVGATEMLRPSFGSIADNPNIHIAQFSRSGATRDGRIPDVTAPGTILSARSSLAADSSSF 445
QY 240 ANNSKXAYWGGTSMATPIVAGNVAOLREHFINKNGITPKPSILKALTAGADVGLGP 299
DB 446 ANNSKXAYWGGTSMATPIVAGNVAOLREHFINKNGITPKPSILKALTAGADVGLGP 505
QY 300 PPSGQRFSLAIYH 433
DB 506 PPSGQRFSLAIYH 641
QY 360 YTLVNDLVLITPAPNOQKYYGNDSEYFYDNNMDGRNNVENVTINAPQSTYITIEVOAYN 419
DB 566 YTLVNDLVLITPAPNOQKYYGNDSEYFYDNNMDGRNNVENVTINAPQSTYITIEVOAYN 625
QY 420 PPSGQRFSLAIYH 433
DB 626 PPSGQRFSLAIYH 641

RESULT 7

US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-8

Query Match 88.2%; Score 1995.5; DB 4; Length 640;
Best Local Similarity 87.6%; Pred. No. 9.1e-146;
Matches 380; Conservative 28; Mismatches 25; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVQNNYGLYGQGOVVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 60
DB 207 NDVARGIVKADVQNNYGLYGQGOVVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 266
QY 61 NADDPNGHGHVAGSVLGN-ALNKGAPOANLVFQSIIMDSGGGLGGLPSNLNTLFSQAWN 119
DB 267 NADDPNGHGHVAGSVLGN-ALNKGAPOANLVFQSIIMDSGGGLGGLPSNLNTLFSQAWN 326
QY 120 AGARIHNSWGA PYNAGYATANSROVDEYVRNNDMTVLFAAGNPGNSGTISAPGTAKNAI 179
DB 327 AGARIHNSWGA PYNAGYATANSROVDEYVRNNDMTVLFAAGNPGNSGTISAPGTAKNAI 386
QY 180 TVGATENYRPSFGSIADNPNIHIAQFSRSGATRDGRIPDVTAPGTILSARSSLAADSSSF 239
DB 387 TVGATEMLRPSFGSIADNPNIHIAQFSRSGATRDGRIPDVTAPGTILSARSSLAADSSSF 446
QY 240 ANNSKXAYWGGTSMATPIVAGNVAOLREHFINKNGITPKPSILKALTAGADVGLGP 299
DB 447 ANNSKXAYWGGTSMATPIVAGNVAOLREHFINKNGITPKPSILKALTAGADVGLGP 506
QY 300 PPSGQRFSLAIYH 433

```
Db 507 PNGNGQWGRVLTLDKSLNVAAYVNESSLSLSQKATYSFATACKPKISLVSDAPASTTA 566
QY 360 STTLVNDLDLVITAPNGQKTYVNDPSYPYDNNWGRNVEVFINAPSGTYTTEVQAYN 419
Db 567 SYTLVNDLDLVITAPNGQTYVNDPFTSPYDNNWGRNVEVFINAPSGTYTTEVQAYN 626
QY 420 VPSGQRFSLATVH 433
Db 627 VPVGPQFSLATVH 640

RESULT 8
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6

Query Match 88.1%; Score 1994.5; DB 4; Length 640;
Best Local Similarity 87.6%; Pred. No. 1.1e-145;
Matches 380; Conservative 28; Mismatches 25; Indels 1; Gaps 1;

QY 1 NVVAGTIVADVAQNNYGLYGQGVVAVADTGLDGRNDSMEAFRGKITALVALGRIN 60
Db 207 NVVAGTIVADVAQSSVGLYGQGVVAVADTGLDGRNDSMEAFRGKITALVALGRIN 266
QY 61 NNDPNHGHTHVASGLVGN-ALNKGMAPQANLVFQSIIDSSGGGLGGLPSNINTLFSQAMN 119
Db 267 NNDPNHGHTHVASGLVGNSTNKGMAPQANLVFQSIIDSSGGGLGGLPSNINTLFSQAMN 326
QY 120 AGARHTNSWGPAPVNGAYTANSRQVDEYVRNNDTVLPAGNEGNSGTTIAPGTAKNAI 179
Db 327 AGARHTNSWGAAYVNGAYTTDSRNDDYRKNDMTILPAAGNEGNGGTTIAPGTAKNAI 386
QY 180 TVGATENYRPSFGSIADNNPHIAQSSRGATRDGRIKPDVTAAPGFIIISASSSLAPDSF 239
Db 387 TVGATENLRFPSFGSIADNNPHIAQSSRGATRDGRIKPDVTAAPGFIIISASSSLAPDSF 446
QY 240 WANVNSKVAVMGCTSMATPIVAGNVAQJREHFIKNRGITPKESLIKAAIAGADVGLGY 299
Db 447 WANVNSKVAVMGCTSMATPIVAGNVAQJREHFVKNRGITPKESLIKAAIAGADVGLGY 506
QY 300 PPSDQGMGVTLDKSLNVAAYNEATALTGOKATYSFQTOAGKPKISLVMTDAPGSTTA 359
Db 507 PPSDQGMGVTLDKSLNVAAYNESSLSQKATYSFATACKPKISLVSDAPASTTA 566
QY 360 SYTLVNDLDLVITAPNGQKTYVNDPSYPYDNNWGRNVEVFINAPSGTYTTEVQAYN 419
Db 567 SYTLVNDLDLVITAPNGQTYVNDPFTSPYDNNWGRNVEVFINAPSGTYTTEVQAYN 626
```

```
QY 420 VPSGQRFSLATVH 433
Db 627 VPVGPQFSLATVH 640

RESULT 9
US-09-509-814A-1
; Sequence 1, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (46)..(46)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (105)..(105)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (130)..(130)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
```

LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (359)..(359)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 86.1%; Score 1948.5; DB 4; Length 639;
Best Local Similarity 86.4%; Pred. No. 3.8e-142;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGGQVVAVADTGLDGRNDSMHEAFRGKITALYALGRN 60
DB NDVARGIVKADVAQSSVGLYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALGRN 265
QY 61 NANDNGHGVHAGSVGN-ALNKGAPOANTVFOGIMSSGGLGSPVNTLFSQAMN 119
DB NANDNGHGVHAGSVGNKNGKGNAPQANLVFQSIMDSXGGLGSPVLTFLFSQAXS 325
QY 120 AGARHTNSWGAIPNGAYTANSRQVDEYVANNMTVLFAAGNBPNSGTISAPGTAKNAI 179
DB AGARHTNSWGAIVNGAYTANSDRVDPYAKNDMTILFAAGNBPNGGTISAPGTAKNAI 385
QY 180 TVGATENYRPSFGIADNPENHIAQFSSRGATRGRIKPDVTAGTFLSRSSLPDSSF 229
DB TVGATENYRPSFGIADNPENHIAQFSSRGATRGRIKPDVTAGTFLSRSSLPDSSF 445
QY 240 WANTSXYAVWGTSVATPIVAGVVAQREHFTKNGITRPSLIALALTAGADVGLGY 299
DB WANTSXYAVWGTSVATPIVAGVVAQREHFTKNGITRPSLIALALTAGADVGLGY 505
QY 300 PSGDQWGRVTLDKSLNVAVNEATLTTQKATYFQIQAGPLKLSLWTTAPSGTTA 359
DB PSGDQWGRVTLDKSLNVAVNEATLTTQKATYFQIQAGPLKLSLWTTAPSGTTA 565
QY 360 SYLVNDLDVITPAENQKTVGNDFSYPYDNNMDGRNNEVFINAPQGTYYIEVQAVN 419
DB SYLVNDLDVITPAENQKTVGNDFSYPYDNNMDGRNNEVFINAPQGTYYIEVQAVN 625
QY 420 VPSPGQRFSLAIYH 433
DB VPSPGQRFSLAIYH 639

RESULT 10
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION: MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHISA
APPLICANT: KUBOTA, HIROKI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEIWA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509, 814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08


```
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
```

```
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2
```

Query Match 86.1%; Score 1948.5; DB 4; length 640;
Best Local Similarity 86.4%; Pred. No. 3.8e-142;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 1 NDVARGIKADVAQNNYGVGGGVAVADTGLDTGNDSSMHEAFRGKTALYALGRTN 60
DB 207 NDVARGIKADVAQSSYGLVGQGVAVADTGLDTGNDSSMHEAFRGKTALYALGRTN 266

```
QY 61 NNDPNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIIMDSGGIGGIPSNLTLLFSCAMN 119
DB 267 NNDPNGHGHVAGSVLGNXNKGMAPOANLVFQSIIMDSXGIGGIPSNLTLLFSCAXS 326
QY 120 AGARIHNSWGAAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNNGTISAPGTAKNAI 179
DB 327 AGARIHNSWGAAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNNGTISAPGTAKNAI 386
QY 180 TVGATENYRPSFGSIANPNHIAOFSSRGATRDGRIPDYTAPOGTILSARSSLAPDSSE 239
DB 387 TVGATENYRPSFGSIANPNHIAOFSSRGATRDGRIPDYTAPOGTILSARSSLAPDSSE 446
QY 240 WANYSKYAVWGGTSMATPIVAGNVAQLREHFVKNRGTIPKPSLIRKALILAGATDVGLGY 299
DB 447 WANYSKYAVWGGTSMATPIVAGNVAQLREHFVKNRGTIPKPSLIRKALILAGATDVGLGY 506
QY 300 PSGDQGMGRVTLDKSLNVAAYNEATALTTOQKATYSEFQAGKPLKISLWTDAPGSTTA 359
DB 507 PNGQGMGRVTLDKSLNVAAYNESXLSLSTQKATYFTATAGKPLKISLWSDAPASTTA 566
QY 360 SYLVNLDLIVITAPNGCKTVGNDPSYEDNNDGRNNEVETINAPOGSTYITTEVOAYN 419
DB 567 SYLVNLDLIVITAPNGCKTVGNDPSYEDNNDGRNNEVETINAPOGSTYITTEVOAYN 626
QY 420 VPSGQGFSLATVH 433
DB 627 VPSGQGFSLATVH 640

RESULT 11
US-09-512-251A-10
; Sequence 10, Application US/09512251A
; Patent No. 655315
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349.204-US
; CURRENT APPLICATION NUMBER: US/09/512.251A
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-512-251A-10
```

```
Query Match 66.9%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 5.2e-109;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNNGVGLGQGVAVADTGLDGTGRNDSMHEAFRGKITALTALYAGRTN 60
DB 29 NDVARGIVKADVAQNNNGVGLGQGVAVADTGLDGTGRNDSMHEAFRGKITALTALYAGRTN 88
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGIGGIPSNLTLLFSCAMNA 120
DB 89 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGIGGIPSNLTLLFSCAYSA 148
QY 121 GARIHNSWGAAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNNGTISAPGTAKNAI 180
DB 149 GARIHNSWGAAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNNGTISAPGTAKNAI 208
QY 181 VGTENYRPSFGSIANPNHIAOFSSRGATRDGRIPDYTAPOGTILSARSSLAPDSSE 240
DB 209 VGTENYRPSFGSIANPNHIAOFSSRGATRDGRIPDYTAPOGTILSARSSLAPDSSE 268
QY 241 ANYNSKYAVWGGTSMATPIVAGNVAQLREHFVKNRGTIPKPSLIRKALILAGATDVGLGP 300
DB 269 ANYNSKYAVWGGTSMATPIVAGNVAQLREHFVKNRGTIPKPSLIRKALILAGADVGLGFP 328
```

```
QY 301 SGDQGMGRVTLDKSLNV 317
DB 329 NGNGQGMGRVTLDKSLNV 345

RESULT 12
US-09-515-150A-10
; Sequence 10, Application US/09515150A
; Patent No. 6558938
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5348.204-US
; CURRENT APPLICATION NUMBER: US/09/515.150A
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-515-150A-10
```

```
Query Match 66.9%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 5.2e-109;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNNGVGLGQGVAVADTGLDGTGRNDSMHEAFRGKITALTALYAGRTN 60
DB 29 NDVARGIVKADVAQNNNGVGLGQGVAVADTGLDGTGRNDSMHEAFRGKITALTALYAGRTN 88
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGIGGIPSNLTLLFSCAMNA 120
DB 89 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGIGGIPSNLTLLFSCAYSA 148
QY 121 GARIHNSWGAAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNNGTISAPGTAKNAI 180
DB 149 GARIHNSWGAAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNNGTISAPGTAKNAI 208
QY 181 VGTENYRPSFGSIANPNHIAOFSSRGATRDGRIPDYTAPOGTILSARSSLAPDSSE 240
DB 209 VGTENYRPSFGSIANPNHIAOFSSRGATRDGRIPDYTAPOGTILSARSSLAPDSSE 268
QY 241 ANYNSKYAVWGGTSMATPIVAGNVAQLREHFVKNRGTIPKPSLIRKALILAGATDVGLGP 300
DB 269 ANYNSKYAVWGGTSMATPIVAGNVAQLREHFVKNRGTIPKPSLIRKALILAGADVGLGFP 328
QY 301 SGDQGMGRVTLDKSLNV 317
DB 329 NGNGQGMGRVTLDKSLNV 345

RESULT 13
US-09-196-281-13
; Sequence 13, Application US/09196281A
; Patent No. 6605458
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/09/196.281A
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 1332/97
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
```

```

; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-196-281-13

Query Match
Best Local Similarity 66.9%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 5.2e-109;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADYKANNNGYLGQGVVAVADTGLDGTGRDSSMHEAFRKTITALLYALGRTN 60
DB 29 NDVARGIVADYKANNNGYLGQGVVAVADTGLDGTGRDSSMHEAFRKTITALLYALGRTN 88
QY 61 NADENGHGTHTVAGAVLGNALNKGMAPQANLVFQSIMDSGGGLGSPNLTILFSGAANA 120
DB 89 NADENGHGTHTVAGAVLGNALNKGMAPQANLVFQSIMDSGGGLGSPNLTILFSGAANA 148
QY 121 GARHTNSGAPVNGAYTANSRQVDEYRANNDMTVLPAAAGNEGPNSTISAPGTAKNAIT 180
DB 149 GARHTNSGAPVNGAYTANSRQVDEYRANNDMTVLPAAAGNEGPNSTISAPGTAKNAIT 208
QY 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGR:KPYTAAGTFTLSARSLSAPDSFW 240
DB 209 VBATENLRPSFGSYADNPNHIAQSSRGATRDGR:KPYTAAGTFTLSARSLSAPDSFW 268
QY 241 ANYNSKRYAMGTSMTPTIVAGNVAQLREHFIKRNKGTIPKPSLIIKALIIAGATDVLGYP 300
DB 269 ANHDSKYAMGTSMTPTIVAGNVAQLREHFIKRNKGTIPKPSLIIKALIIAGATDVLGYP 328
QY 301 SGDOGGRVTLDKSLNV 317
DB 329 NGNQGGRVTLDKSLNV 345

RESULT 14
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION NUMBER: JP 322285/1995
; APPLICATION DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618

```

```

; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-818B-1

Query Match
Best Local Similarity 19.6%; Score 443; DB 3; Length 659;
Best Local Similarity 31.1%; Pred. No. 4.8e-26;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;

QY 8 VKADVAKNNYGLYGQGVVAVADTGLDGTGRDSSMHEAFRKTITALLY-ALGRTNANDPN 66
DB 145 IGADYVNSLGYDGSQVVAIVDTGIDAN-----HPDLKGVIGWYDAVNGRSTPYDDQ 198
QY 67 GHGTHVAGSYLG---NALNKGMAPQANLVFQSIM--DSGGGLGSPNLTILFSGAANA 120
DB 199 GHGTHVAGSYLGSYNSQYIGVAPAKLYGVKVLGADSSGSVSTIIAGDVWVQKDKX 258
QY 121 GARHTNSGAPVNGAYTANSRQVDEYRANNDMTVLPAAAGNEGPNSTISAPGTAKNAIT 180
DB 259 GIRVINSLSGSSQSGTSLSQAVNNAMDAGIVCVAAGNSGFNTYTGSPALASKVIT 318
QY 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGR:KPYTAAGTFTLSARSLSAPDSFW 240
DB 319 VCA-----VSDNIIASFSSRGPTADGRLEPYVAVROVDIARAS---GTSNG 364
QY 241 ANYNSKRYAMGTSMTPTIVAGNVAQLREHFIKRNKGTIP--KPSLIIKALIIAGATDVG- 296
DB 365 TPINDYTTKAGTSMATPHVSGVGLIIQ---AHPSTPDKVTALIEFDIADIAPEKID 421
QY 297 LGYPSDQGWGRVTLDKSLNVAVNEATALTGQ-----KATSPQTQAGKPLKISLVMT 351
DB 422 IAY---GGRVNVVKA--IKTDYAKLITTSVADKSAITHTFVSGATFTATLYND 474
QY 352 DAGSTTASYTLVNDLVTTAPNGQKYGNDSEPYDNMNGRNVENVFINAPQSGTY 411
DB 475 -----TGSSDIDLVDYDN-----GNEVDYGYTAYY---GFEKVGYNPFRAGT 515
QY 412 TIEVQANVPSGQPSRLAIY 432
DB 516 TVKVVSY--KGANYQVDVY 533

RESULT 15
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-09-445-472-12

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 31, 2004, 16:09:09 ; Search time 33.1412 Seconds

(without alignment)
3418.697 Million cell updates/sec

Title: US-09-985-689A-3

Perfect score: 2263

Sequence: 1 NDVARGIVKADVAGNNYGLY.....EYQATNVSGGQRFSLATVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 26161801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2263	100.0	433	10	US-09-985-689A-3
2	2251	99.5	433	10	US-09-985-689A-5
3	2234	98.7	433	10	US-09-985-689A-4
4	2040	90.1	433	10	US-09-985-689A-7
5	2020.5	89.3	434	10	US-09-985-689A-6
6	2006.5	88.7	434	10	US-09-985-689A-2
7	1994.5	88.1	434	10	US-09-985-689A-1
8	1994.5	88.1	434	10	US-10-385-682-2
9	1514	66.9	345	14	US-10-385-682-1
10	1514	66.9	345	14	US-10-403-105-13
11	443	19.6	659	13	US-10-090-624-12
12	422.5	18.7	412	13	US-10-090-624-11
13	422.5	18.7	522	13	US-10-090-624-4
14	422.5	18.7	654	13	US-10-090-624-16
15	346.5	15.3	1237	14	US-10-314-657-4

16	345.5	15.3	1079	14	US-10-112-488-39	Sequence 39, Appl
17	329	14.5	1208	14	US-10-156-761-13251	Sequence 13251, A
18	327	14.4	1139	14	US-10-156-761-10856	Sequence 10856, A
19	305	13.5	1398	13	US-10-090-624-6	Sequence 6, Appl
20	296.5	13.1	580	10	US-09-927-827-55	Sequence 55, Appl
21	295	13.0	519	15	US-10-084-846A-114	Sequence 114, App
22	295	13.0	19725	15	US-10-084-846A-4	Sequence 4, Appl
23	282.5	12.5	1101	14	US-10-156-761-12934	Sequence 12934, A
24	281.5	12.4	271	10	US-10-344-231-3	Sequence 3, Appl
25	276	12.2	271	10	US-09-813-408-2	Sequence 2, Appl
26	271	12.0	271	14	US-10-242-549-56	Sequence 56, Appl
27	270	11.9	271	14	US-10-242-549-54	Sequence 54, Appl
28	270	11.9	271	14	US-10-242-549-60	Sequence 60, Appl
29	269	11.9	269	8	US-08-322-578-10	Sequence 10, Appl
30	269	11.9	269	9	US-09-837-335-16	Sequence 16, Appl
31	269	11.9	269	9	US-09-060-854B-6	Sequence 6, Appl
32	269	11.9	269	9	US-09-975-139-1	Sequence 1, Appl
33	269	11.9	269	9	US-09-976-114-8	Sequence 8, Appl
34	269	11.9	269	10	US-09-736-116-49	Sequence 49, Appl
35	269	11.9	269	13	US-10-075-907-1	Sequence 1, Appl
36	269	11.9	269	13	US-10-075-895-1	Sequence 1, Appl
37	269	11.9	269	14	US-10-033-325-6	Sequence 6, Appl
38	269	11.9	269	14	US-10-209-812-3	Sequence 3, Appl
39	269	11.9	269	14	US-10-104-693-4	Sequence 4, Appl
40	269	11.9	269	14	US-10-228-572-6	Sequence 6, Appl
41	269	11.9	269	14	US-10-324-152-5	Sequence 5, Appl
42	269	11.9	269	14	US-10-242-549-5	Sequence 5, Appl
43	269	11.9	269	14	US-10-313-853-1	Sequence 1, Appl
44	269	11.9	269	14	US-10-336-324-4	Sequence 4, Appl
45	269	11.9	269	14	US-10-336-324-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-985-689A-3
Sequence 3, Application US/09985689A
Publication No. US20030022351A1
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TOSYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKETI, KATSUISHI
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483USO
CURRENT APPLICATION NUMBER: US/09/985,689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 433
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-3
Query Match 100.0%; Score 2263; DB 10; Length 433;
Best Local Similarity 100.0%; Pred. No. 3.5e-191;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 NDVARGIVKADVAGNNYGLGGQGVAVADTGDGTGNDSSMGEAFRGKTTALYALGRN 60
1 NDVARGIVKADVAGNNYGLGGQGVAVADTGDGTGNDSSMGEAFRGKTTALYALGRN 60
61 NADPNHGHTVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLITLFSQAWNA 120

```

Db      61  NADPNHGHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Qy      121  GAIHTNSWGA PYNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121  GAIHTNSWGA PYNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Qy      181  VGATENYRPSFGSIADPNPHIAQFSSRGATRDRIKPDVTAPGTFTLSARSSIAPOSSFW 240
Db      181  VGATENYRPSFGSIADPNPHIAQFSSRGATRDRIKPDVTAPGTFTLSARSSIAPOSSFW 240
Qy      241  ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPSPSLIKALLAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPSPSLIKALLAGATDVGLGYP 300
Qy      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFQOAGKPKLSLWTDAPGSTTAS 360
Db      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFQOAGKPKLSLWTDAPGSTTAS 360
Qy      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNEVVFINAPQSGTYTIEVOAYNV 420
Db      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNEVVFINAPQSGTYTIEVOAYNV 420
Qy      421  PSGPQRFSLAIHV 433
Db      421  PSGPQRFSLAIHV 433

```

RESULT 2

```

US-09-985-689A-5
; Sequence 5, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-5

```

```

Query Match      99.5%; Score 2251; DB 10; Length 433;
Best Local Similarity 99.3%; Pred. No. 4,1e-190;
Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy      1  NDVARGIVKADVQNNYGLYGQGVVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Db      1  NDVARGIVKADVQNNYGLYGQGVVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Qy      61  NADPNHGHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Db      61  NADPNHGHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Qy      121  GAIHTNSWGA PYNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121  GAIHTNSWGA PYNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180

```

```

Qy      181  VGATENYRPSFGSIADPNPHIAQFSSRGATRDRIKPDVTAPGTFTLSARSSIAPOSSFW 240
Db      181  VGATENYRPSFGSIADPNPHIAQFSSRGATRDRIKPDVTAPGTFTLSARSSIAPOSSFW 240
Qy      241  ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPSPSLIKALLAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPSPSLIKALLAGATDVGLGYP 300
Qy      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFQOAGKPKLSLWTDAPGSTTAS 360
Db      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFQOAGKPKLSLWTDAPGSTTAS 360
Qy      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNEVVFINAPQSGTYTIEVOAYNV 420
Db      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNEVVFINAPQSGTYTIEVOAYNV 420
Qy      421  PSGPQRFSLAIHV 433
Db      421  PSGPQRFSLAIHV 433

```

RESULT 3

```

US-09-985-689A-4
; Sequence 4, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-4

```

```

Query Match      98.7%; Score 2234; DB 10; Length 433;
Best Local Similarity 98.6%; Pred. No. 1,3e-188;
Matches 427; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy      1  NDVARGIVKADVQNNYGLYGQGVVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Db      1  NDVARGIVKADVQNNYGLYGQGVVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Qy      61  NADPNHGHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Db      61  NADPNHGHGTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Qy      121  GAIHTNSWGA PYNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121  GAIHTNSWGA PYNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Qy      181  VGATENYRPSFGSIADPNPHIAQFSSRGATRDRIKPDVTAPGTFTLSARSSIAPOSSFW 240
Db      181  VGATENYRPSFGSIADPNPHIAQFSSRGATRDRIKPDVTAPGTFTLSARSSIAPOSSFW 240
Qy      241  ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPSPSLIKALLAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMAPIVAGNVAQLREHFINKRGITPSPSLIKALLAGATDVGLGYP 300

```

QY 301 SGGDQMGRTYLDKSLNVAVNEATALTGGOKATYSFOTQAGKPLKISLWMTDAPGSTTAS 360
 Db 301 NDDQMGRTYLDKSLNVAVNEATALTGGOKATYSFOTQAGKPLKISLWMTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKYVNDFSYPYDNNDGNNVENVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGQKYVNDFSYPYDNNDGNNVENVFINAPQSGTYTIEVOAYNV 420
 QY 421 PGGPQRFSLAIYH 433
 Db 421 PGGPQRFSLAIYH 433

RESULT 4
 US-09-985-689A-7
 ; Sequence 7, Application US/09985689A
 ; Publication No. US20030022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483US0
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 433
 ; TYPE: PRT
 ; ORGANISM: Bacillus sp.
 ; US-09-985-689A-7

Query Match 90.1%; Score 2040; DB 10; Length 433;
 Best Local Similarity 88.9%; Pred. No. 1,8e-171;
 Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;

QY 1 NDVARGIVADVAONNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 Db 1 NDVARGIVADVAONNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 QY 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 120
 Db 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 120
 QY 121 GARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSGTISAPGTAKNAIT 180
 Db 121 GARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 240
 Db 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 240
 QY 241 ANYSKYAYMGTSNATPIVAGNVAAQLREHFKNRGITPKPSLAKAALIAGATDVGGLGP 300
 Db 241 ANYSKYAYMGTSNATPIVAGNVAAQLREHFKNRGITPKPSLAKAALIAGATDVGGLGP 300
 QY 301 SGGDQMGRTYLDKSLNVAVNEATALTGGOKATYSFOTQAGKPLKISLWMTDAPGSTTAS 360
 Db 301 SGGDQMGRTYLDKSLNVAVNEATALTGGOKATYSFOTQAGKPLKISLWMTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKYVNDFSYPYDNNDGNNVENVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGQKYVNDFSYPYDNNDGNNVENVFINAPQSGTYTIEVOAYNV 420

Db 361 LTLVNDLDELVTAPNGKTYVNDFTAPYDNNDGNNVENVFINAPQSGTYTIEVOAYNV 420
 QY 421 PGGPQRFSLAIYH 433
 Db 421 PGGPQRFSLAIYH 433

RESULT 5
 US-09-985-689A-6
 ; Sequence 6, Application US/09985689A
 ; Publication No. US20030022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483US0
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 434
 ; TYPE: PRT
 ; ORGANISM: Bacillus sp.
 ; US-09-985-689A-6

Query Match 89.3%; Score 2020.5; DB 10; Length 434;
 Best Local Similarity 88.2%; Pred. No. 9.3e-170;
 Matches 383; Conservative 28; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIVADVAONNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 Db 1 NDVARGIVADVAONNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 QY 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 119
 Db 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 120
 QY 120 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSGTISAPGTAKNAI 179
 Db 120 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSGTISAPGTAKNAI 180
 QY 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 239
 Db 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 240
 QY 240 MANTYKAYAMGTSNATPIVAGNVAAQLREHFKNRGITPKPSLAKAALIAGATDVGGLGP 299
 Db 240 MANTYKAYAMGTSNATPIVAGNVAAQLREHFKNRGITPKPSLAKAALIAGATDVGGLGP 300
 QY 300 SGGDQMGRTYLDKSLNVAVNEATALTGGOKATYSFOTQAGKPLKISLWMTDAPGSTTAS 359
 Db 300 SGGDQMGRTYLDKSLNVAVNEATALTGGOKATYSFOTQAGKPLKISLWMTDAPGSTTAS 360
 QY 360 SYTLVNDLDELVTAPNGQKYVNDFSYPYDNNDGNNVENVFINAPQSGTYTIEVOAYNV 419
 Db 360 SYTLVNDLDELVTAPNGQKYVNDFSYPYDNNDGNNVENVFINAPQSGTYTIEVOAYNV 420
 QY 420 VPGPQRFSLAIYH 433
 Db 420 VPGPQRFSLAIYH 434

RESULT 6

US-09-985-689a-2
 ; Sequence 2, Application US/09985689a
 ; Publication No. US2003022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SAITO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483US6
 ; CURRENT APPLICATION NUMBER: US/09/985, 689A
 ; PRIORITY FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 434
 ; TYPE: PRT
 ; ORGANISM: Bacillus sp.
 US-09-985-689a-2

Query Match 88.7%; Score 2006.5; DB 10; Length 434;
 Best Local Similarity 88.2%; Pred. No. 1.6e-168;
 Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
 DB 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
 QY 61 NANDPNHGHTHVAGSVLGN ALNKGMAPQANLVFQSIIMSSGGLGGLPSNLTLFSGQAN 119
 DB 61 NANDPNHGHTHVAGSVLGNATNKGMAPQANLVFQSIIMSSGGLGGLPSNLTLFSGQAS 120
 QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPRSGTISAPGTAKNAI 179
 DB 121 AGARIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPRSGTISAPGTAKNAI 180
 QY 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLABDSF 239
 DB 181 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLABDSF 240
 QY 240 WANNYSKYAYMGSTSMATPIVAGNVAQLREHFIKNGGITPKPSLKAALIAAGTVDGLGY 299
 DB 241 WANNYSKYAYMGSTSMATPIVAGNVAQLREHFIKNGGITPKPSLKAALIAAGTVDGLGY 300
 QY 300 PPSGQGWGRVTLDKSLNVAAYNEATLTTGQKATYSFOTQAGKPLKISLWTDAPGSTTA 359
 DB 301 PPSGQGWGRVTLDKSLNVAAYNEATLTTGQKATYSFOTQAGKPLKISLWTDAPGSTTA 360
 QY 360 SYTLVNDLDTVITAPNGOKTVGNDFSYPYDNNMDGNNVENVFINAPQSGTYTIEVOAYN 419
 DB 361 SYTLVNDLDTVITAPNGOKTVGNDFSYPYDNNMDGNNVENVFINAPQSGTYTIEVOAYN 420
 QY 420 VPSGQGRFSLAIYH 433
 DB 421 VPSGQGRFSLAIYH 434

RESULT 7

US-09-985-689a-1
 ; Sequence 1, Application US/09985689a
 ; Publication No. US2003022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI

APPLICANT: KAGEYAMA, YASUSHI
 APPLICANT: SAITO, TSUYOSHI
 APPLICANT: ARAKI, HIROYUKI
 APPLICANT: SUMITOMO, NOBUYUKI
 APPLICANT: OKUDA, MITSUYOSHI
 APPLICANT: SAEKI, KATSUHIISA
 TITLE OF INVENTION: Alkaline proteases
 FILE REFERENCE: 215483US6
 CURRENT APPLICATION NUMBER: US/09/985, 689A
 PRIORITY FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: JP P2000-355166
 PRIOR FILING DATE: 2000-11-22
 PRIOR APPLICATION NUMBER: JP P2001-114048
 PRIOR FILING DATE: 2001-04-12
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 434
 TYPE: PRT
 ORGANISM: Bacillus sp.
 US-09-985-689a-1

Query Match 88.1%; Score 1994.5; DB 10; Length 434;
 Best Local Similarity 87.6%; Pred. No. 1.6e-167;
 Matches 380; Conservative 28; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
 DB 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
 QY 61 NANDPNHGHTHVAGSVLGN ALNKGMAPQANLVFQSIIMSSGGLGGLPSNLTLFSGQAN 119
 DB 61 NANDPNHGHTHVAGSVLGNSTNKGMAPQANLVFQSIIMSSGGLGGLPSNLTLFSGQAS 120
 QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPRSGTISAPGTAKNAI 179
 DB 121 AGARIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPRSGTISAPGTAKNAI 180
 QY 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLABDSF 239
 DB 181 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLABDSF 240
 QY 240 WANNYSKYAYMGSTSMATPIVAGNVAQLREHFIKNGGITPKPSLKAALIAAGTVDGLGY 299
 DB 241 WANNYSKYAYMGSTSMATPIVAGNVAQLREHFIKNGGITPKPSLKAALIAAGTVDGLGY 300
 QY 300 PPSGQGWGRVTLDKSLNVAAYNEATLTTGQKATYSFOTQAGKPLKISLWTDAPGSTTA 359
 DB 301 PPSGQGWGRVTLDKSLNVAAYNEATLTTGQKATYSFOTQAGKPLKISLWTDAPGSTTA 360
 QY 360 SYTLVNDLDTVITAPNGOKTVGNDFSYPYDNNMDGNNVENVFINAPQSGTYTIEVOAYN 419
 DB 361 SYTLVNDLDTVITAPNGOKTVGNDFSYPYDNNMDGNNVENVFINAPQSGTYTIEVOAYN 420
 QY 420 VPSGQGRFSLAIYH 433
 DB 421 VPSGQGRFSLAIYH 434

RESULT 8

US-10-385-662-2
 ; Sequence 2, Application US/10385662
 ; Publication No. US20040002432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SAITO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: IZAMA, YOSHIFUMI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; APPLICANT: KOBAYASHI, TOHRU
 ; APPLICANT: NOMURA, MASAFUMI
 ; TITLE OF INVENTION: Alkaline protease


```
Db 149 GARHTNSWGAFFVGTDSRVVDYVRKNDWTLFAAGNPGSGTISAPGTAKNAIT 208
Qy 181 VGATENYRPSFGSIADNPENHIAOFSSRGATROGRIRPDVTAGFTILSARSLADSSFW 240
Db 209 VGATENYRPSFGSIADNPENHIAOFSSRGATROGRIRPDVTAGFTILSARSLADSSFW 268
Qy 241 ANNSKXAYWGTSMATPIVAGNVAAQLREHFINKRGITKPSILKALJAGATDVGLYP 300
Db 269 ANNSKXAYWGTSMATPIVAGNVAAQLREHFINKRGITKPSILKALJAGATDVGLYP 328
Qy 301 SGDOGWRVTLDSLVN 317
Db 329 NSNGQWGRVTLDSLVN 345

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PR
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 19.6%; Score 443; DB 13; Length 659;
Best Local Similarity 31.1%; Pred. No. 3,7e-30;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;

Qy 8 VKADVAQNNGVGLGQGVAAVADTGLDGRNDSKHEAFRGKITALY-ALGRTNNANPN 66
Db 145 TGAADTYWNSLGYSGVVAIVDTGIDAN-----HPDLKGKVIQWYDANGRSTPYDQ 198
Qy 67 GHGTHVAGSVLG-----NALNKGAPQANLVFQSIW--DSSGGLGSLPSNLTLSFQAWNA 120
Db 199 GHGTHVAGIAGTGVNSQYIGVAPGAKLVGVKVLGADGSGSVITIIADVWVQNKXK 258
Qy 121 GARIHNSMGAPVNGAVTANSRQVDEVYRNNDMTYLPAAQNGEPNSGTISAGTAKNAIT 180
Db 259 GARIHNSMGAPVNGAVTANSRQVDEVYRNNDMTYLPAAQNGEPNSGTISAGTAKNAIT 318
Qy 181 VGATENYRPSFGSIADNPENHIAOFSSRGATROGRIRPDVTAGFTILSARSLADSSFW 240
Db 319 VGATENYRPSFGSIADNPENHIAOFSSRGATROGRIRPDVTAGFTILSARSLADSSFW 364
Qy 241 ANNSKXAYWGTSMATPIVAGNVAAQLREHFINKRGITP---KPSILKALJAGATDVG- 296
Db 365 TPINDYTTKASGTSMTAPVAGNVAAQLREHFINKRGITP---KPSILKALJAGATDVG- 421
Qy 297 LGYPSDQWGRVTLDSLVNVAVNEATALTTCQ-----KATYSFQTAGKPLKISLVWT 351
Db 422 LGYPSDQWGRVTLDSLVNVAVNEATALTTCQ-----KATYSFQTAGKPLKISLVWT 474
Qy 352 DAGGSTTATYLVNDLDTVITAPNGQKYVGNDSFYDNNMNGRNNVENVFINAQSGTY 411
Db 475 DAGGSTTATYLVNDLDTVITAPNGQKYVGNDSFYDNNMNGRNNVENVFINAQSGTY 515

US-10-090-624-12
; Sequence 1, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PR
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.7%; Score 422.5; DB 13; Length 412;
Best Local Similarity 31.4%; Pred. No. 1,2e-28;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

Qy 16 NYGAYGQGVAAVADTGLDGRNDSKHEAFRGKITALYALGRTNNAN-----DENGHG 69
Db 20 NLGYDGSIGTITIGITIDID-----ASHPDLQKV-----IGWVDVNGRSYDDHGHG 68
Qy 70 THVAGSVLG-----NALNKGAPQANLVFQSIW--DSSGGLGSLPSNLTLSFQAWNA 122
Db 69 THVAGSVLG-----NALNKGAPQANLVFQSIW--DSSGGLGSLPSNLTLSFQAWNA 128
Qy 123 RHHTNSMGAPVNGAVTANSRQVDEVYRNNDMTYLPAAQNGEPNSGTISAGTAKNAIT 182
Db 123 RHHTNSMGAPVNGAVTANSRQVDEVYRNNDMTYLPAAQNGEPNSGTISAGTAKNAIT 188
Qy 183 ATENYRPSFGSIADNPENHIAOFSSRGATROGRIRPDVTAGFTILSARSLADSSFWAN 242
Db 189 AVDKY-----DVITFSRSGPTADGRKPEVVAFGNIIAFAAS---GTSMGQP 234
Qy 243 YNSKXAYWGTSMATPIVAGNVAAQLREHFINKRGITP---KPSILKALJAGATDVG- 298
Db 235 INDYTTAPGTSMTAPVAGNVAAQLREHFINKRGITP---KPSILKALJAGATDVG- 353
Qy 299 YPSGDQWGRVTLDSLVNVAVNEATALTTCQ-----KATYSFQTAGKPLKISLVWT 353
Db 299 YPSGDQWGRVTLDSLVNVAVNEATALTTCQ-----KATYSFQTAGKPLKISLVWT 353
Qy 292 Y-----CAGRVNAYKAIN--YDNYAKLVFTGVYANKSGCTHGVISGASFTVATLYWDN 344
Db 354 Y-----CAGRVNAYKAIN--YDNYAKLVFTGVYANKSGCTHGVISGASFTVATLYWDN 344
Qy 345 N-----SDLDLTVPNNGNV---DYSY-----TAYYGFKRVGYNFTDGTWT 385
Db 414 EVQAVNPSGPFQSFSLAIV 432
Qy 386 KVSAY-----SGSANYQVDVV 401

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
```

```
Qy 412 TIEQAVNPSGPFQSFSLAIV 432
Db 516 TVKVSY-----KGAANYQVDVV 533

RESULT 12
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PR
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.7%; Score 422.5; DB 13; Length 412;
Best Local Similarity 31.4%; Pred. No. 1,2e-28;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

Qy 16 NYGAYGQGVAAVADTGLDGRNDSKHEAFRGKITALYALGRTNNAN-----DENGHG 69
Db 20 NLGYDGSIGTITIGITIDID-----ASHPDLQKV-----IGWVDVNGRSYDDHGHG 68
Qy 70 THVAGSVLG-----NALNKGAPQANLVFQSIW--DSSGGLGSLPSNLTLSFQAWNA 122
Db 69 THVAGSVLG-----NALNKGAPQANLVFQSIW--DSSGGLGSLPSNLTLSFQAWNA 128
Qy 123 RHHTNSMGAPVNGAVTANSRQVDEVYRNNDMTYLPAAQNGEPNSGTISAGTAKNAIT 182
Db 123 RHHTNSMGAPVNGAVTANSRQVDEVYRNNDMTYLPAAQNGEPNSGTISAGTAKNAIT 188
Qy 183 ATENYRPSFGSIADNPENHIAOFSSRGATROGRIRPDVTAGFTILSARSLADSSFWAN 242
Db 189 AVDKY-----DVITFSRSGPTADGRKPEVVAFGNIIAFAAS---GTSMGQP 234
Qy 243 YNSKXAYWGTSMATPIVAGNVAAQLREHFINKRGITP---KPSILKALJAGATDVG- 298
Db 235 INDYTTAPGTSMTAPVAGNVAAQLREHFINKRGITP---KPSILKALJAGATDVG- 353
Qy 299 YPSGDQWGRVTLDSLVNVAVNEATALTTCQ-----KATYSFQTAGKPLKISLVWT 353
Db 299 YPSGDQWGRVTLDSLVNVAVNEATALTTCQ-----KATYSFQTAGKPLKISLVWT 353
Qy 292 Y-----CAGRVNAYKAIN--YDNYAKLVFTGVYANKSGCTHGVISGASFTVATLYWDN 344
Db 354 Y-----CAGRVNAYKAIN--YDNYAKLVFTGVYANKSGCTHGVISGASFTVATLYWDN 344
Qy 345 N-----SDLDLTVPNNGNV---DYSY-----TAYYGFKRVGYNFTDGTWT 385
Db 414 EVQAVNPSGPFQSFSLAIV 432
Qy 386 KVSAY-----SGSANYQVDVV 401

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
```

APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: SHIMOJO, Tomoko
 APPLICANT: ASADA, Kiyozo
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 FILE REFERENCE: TAKAKURA-6
 CURRENT APPLICATION NUMBER: US/10/090,624
 CURRENT FILING DATE: 2002-03-06
 PRIOR APPLICATION NUMBER: 09/445,472
 PRIOR FILING DATE: 1999-12-06
 PRIOR APPLICATION NUMBER: 151969/1997
 PRIOR FILING DATE: 1997-06-10
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 4
 LENGTH: 522
 TYPE: PR
 ORGANISM: Pyrococcus furiosus
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (428)..(428)
 OTHER INFORMATION: Xaa at position 428 is Gly or Val.
 US-10-090-624-4

Query Match 18.7%; Score 422.5; DB 13; Length 522;
 Best Local Similarity 31.4%; Pred. No. 1,7e-28;
 Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGGQVAVADTGLDGRNDSMEAFRGKITLALGRTNNAN-----DPNGHG 69
 DB 20 NLGYDSGGITIGITIDGID-----ASHPDLQCKV-----IGWDFVNGRSYFYDDHGHG 68
 QY 70 THVAGSVLG-----NALNKGMAPQANLVFQSIW--DSSGGLGGLPSNLNTLFSQAMNAGA 122
 DB 69 THVASIAAGTGAASNGKYKMAPGAKLAGIKVLGADSGSISITIKGVEMAVDNKDKYGI 128
 QY 123 RIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFPAAGNEGPNSGTISAAGTAKAATVVG 182
 DB 129 KVINLSIGSSQSDGTALSOAVNAAMDAGLVVVAAGNSGPKYKTIQSPAASKYITVG 188
 QY 183 ATENYRSPFGSIADNPNHIAQFSSRGATRDGRKPDVTAPGFTILSARSSLPDSSFWAN 242
 DB 189 AVDKY-----DVITSSFSRGPTADGRKPEVVAAGNMIIARAS---GTSWGP 234
 QY 243 YNSKYAVMGTSNATPIVAGNVAQLREHFIKNGITP---KPSLIKALIAGATDVG-LG 298
 DB 235 INDYTTAPGTSNATPVAAGIALLLQ---AHPSWTPDKVKTALIEADIVKPELADIA 291
 QY 299 YPSGDQMGRTVLDKSLNVAVNEATALTTGOKA-----TYSQTQAGKPLKISLVWTD 353
 DB 292 Y-----GAGRNVAKAIN--YDNVAKLVFTGYANKSQTHTQVIGASFTVATLWMDA 344
 QY 354 PGSTTASYTLVNDLDTLITAPNGQKYVNDPSYDNNMDGRNNVENVFTMAPQSGTYTI 413
 DB 345 N-----SDLDLYLDPNGNQV---DYSY-----TAYYGFEXVGYNPTDGTWTI 385
 QY 414 EVQAYNVPSGQRFSLAIV 432
 DB 386 KVVSY---SGSANYQVDV 401
 RESULT 14
 US-10-090-624-16
 Sequence 16, Application US/10090624
 Publication No. US20020132335A1
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: SHIMOJO, Tomoko
 APPLICANT: ASADA, Kiyozo
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

FILE REFERENCE: TAKAKURA-6
 CURRENT APPLICATION NUMBER: US/10/090,624
 CURRENT FILING DATE: 2002-03-06
 PRIOR APPLICATION NUMBER: 09/445,472
 PRIOR FILING DATE: 1999-12-06
 PRIOR APPLICATION NUMBER: 151969/1997
 PRIOR FILING DATE: 1997-06-10
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 16
 LENGTH: 654
 TYPE: PR
 ORGANISM: Pyrococcus furiosus
 US-10-090-624-16

Query Match 18.7%; Score 422.5; DB 13; Length 654;
 Best Local Similarity 31.4%; Pred. No. 2,4e-28;
 Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGGQVAVADTGLDGRNDSMEAFRGKITLALGRTNNAN-----DPNGHG 69
 DB 152 NLGYDSGGITIGITIDGID-----ASHPDLQCKV-----IGWDFVNGRSYFYDDHGHG 200
 QY 70 THVAGSVLG-----NALNKGMAPQANLVFQSIW--DSSGGLGGLPSNLNTLFSQAMNAGA 122
 DB 201 THVASIAAGTGAASNGKYKMAPGAKLAGIKVLGADSGSISITIKGVEMAVDNKDKYGI 260
 QY 123 RIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFPAAGNEGPNSGTISAAGTAKAATVVG 182
 DB 261 KVINLSIGSSQSDGTALSOAVNAAMDAGLVVVAAGNSGPKYKTIQSPAASKYITVG 320
 QY 183 ATENYRSPFGSIADNPNHIAQFSSRGATRDGRKPDVTAPGFTILSARSSLPDSSFWAN 242
 DB 321 AVDKY-----DVITSSFSRGPTADGRKPEVVAAGNMIIARAS---GTSWGP 366
 QY 243 YNSKYAVMGTSNATPIVAGNVAQLREHFIKNGITP---KPSLIKALIAGATDVG-LG 298
 DB 367 INDYTTAPGTSNATPVAAGIALLLQ---AHPSWTPDKVKTALIEADIVKPELADIA 423
 QY 299 YPSGDQMGRTVLDKSLNVAVNEATALTTGOKA-----TYSQTQAGKPLKISLVWTD 353
 DB 424 Y-----GAGRNVAKAIN--YDNVAKLVFTGYANKSQTHTQVIGASFTVATLWMDA 476
 QY 354 PGSTTASYTLVNDLDTLITAPNGQKYVNDPSYDNNMDGRNNVENVFTMAPQSGTYTI 413
 DB 477 N-----SDLDLYLDPNGNQV---DYSY-----TAYYGFEXVGYNPTDGTWTI 517
 QY 414 EVQAYNVPSGQRFSLAIV 432
 DB 516 KVVSY---SGSANYQVDV 533

RESULT 15
 US-10-314-657-4
 Sequence 4, Application US/10314657
 Publication No. US20030175888A1
 GENERAL INFORMATION:
 APPLICANT: SHEN, Ben
 APPLICANT: TANG, Gong-li
 TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polypeptide
 TITLE OF INVENTION: Synthesases and Methods of Use
 FILE REFERENCE: 054030-0021
 CURRENT APPLICATION NUMBER: US/10/314,657
 CURRENT FILING DATE: 2002-12-09
 PRIOR APPLICATION NUMBER: PCT/US02/08937
 PRIOR FILING DATE: 2002-03-22
 PRIOR APPLICATION NUMBER: US 60/278,935
 PRIOR FILING DATE: 2001-03-26
 NUMBER OF SEQ ID NOS: 214
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 4
 LENGTH: 1237

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3246 Seconds

(without alignments)
3677.911 Million cell updates/sec

Title: US-09-985-689A-3

Sequence: 1 NDVARGIVADVAQNNGYL.....EVQAVNPDSGQRFSLAIIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:
2: p1r2:
3: p1r3:
4: p1r4:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	502	22.2	1743	2	T18279 multidrug resistan
2	463.5	20.5	1905	2	T18267 intracellular alka
3	344	15.2	444	2	B83891 intracellular alka
4	315.5	13.9	442	2	A69587 intracellular alka
5	312	13.8	806	2	A41341 microbial serine p
6	304.5	13.5	799	2	G83753 subcellular-type pr
7	303	13.4	1398	2	T28159 pyrolysin (EC 3.4.
8	296.5	13.1	380	2	S11890 serine proteinase
9	288	12.7	1345	2	T29090 surface layer-asso
10	269.5	11.9	715	2	UC4908 alkaline serine pr
11	269	11.9	380	2	A49778 high-alkaline seri
12	265.5	11.7	420	1	S23407 subcellulin (EC 3.4
13	265	11.7	534	1	JS0173 alkaline proteins
14	258.5	11.4	419	1	S25835 subcellulin (EC 3.4
15	256.5	11.3	627	2	D75393 serine proteinase,
16	256	11.3	757	2	C84120 subcellulin-type pr
17	255	11.3	401	2	A72647 serine proteinase
18	254.5	11.2	1331	2	A72647 probable surface 1
19	251	11.1	488	2	A1930 proteinase (import
20	248.5	11.0	1167	1	A35066 streptococcal Csa
21	246	10.9	894	2	F69730 cell wall-associat
22	243.5	10.8	378	2	A33973 high-alkaline seri
23	243	10.7	382	2	I39780 subcellulin (EC 3.4
24	243	10.7	379	1	A35742 aqualysin (EC 3.4.
25	242	10.7	379	1	SUBSCL subcellulin (EC 3.4
26	239.5	10.6	321	1	S27501 alkaline proteins
27	238	10.5	384	1	J24802 alkaline proteins
28	236.5	10.5	402	1	JU0332 alkaline proteins
29	233.5	10.3	319	2	I39866 microbial serine p

30	229.5	10.1	519	2	S71451 halolysin R4 (EC 3
31	228	10.1	1052	2	T17093 intraluminal subtl
32	227.5	10.1	381	1	SUBSS subcellulin (EC 3.4
33	227.5	10.1	381	2	J01487 subcellulin (EC 3.4
34	227.5	10.1	381	2	JH0778 subcellulin (EC 3.4
35	226.5	10.0	381	1	SUBSI subcellulin (EC 3.4
36	226.5	10.0	905	2	F82734 serine proteinase
37	226.5	10.0	1433	1	A36734 bacillopeptidase F
38	225	9.9	382	1	SUBSN subcellulin (EC 3.4
39	224.5	9.9	272	2	A23624 subcellulin (EC 3.4
40	224.5	9.9	323	2	I39667 microbial serine p
41	224	9.9	274	1	SUBSD subcellulin (EC 3.4
42	224	9.9	601	2	JC4576 serine proteinase
43	223.5	9.9	279	1	SUMTV serine proteinase
44	223	9.9	275	1	JC1085 subcellulin (EC 3.4
45	223	9.9	613	2	S73976 hypothetical prote

ALIGNMENTS

RESULT 1	
T18279	multidrug resistance, transport protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum	
C:Date: 15-Oct-1999	#sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18279	
R:Shaulsky, G.; Loomis, W.F.	
Submitted to the EMBL Data Library, June 1996	
Reference number: Z18855	
A:Accession: T18279	
A:Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: mRNA	
A:Residues: 1-1743 <SHA>	
A:Cross-references: EMBL:U60086; NID:g1399914; PID:g1399915; PIDN:AA03331.1	
A:Genetics:	
A:Gene: tagc	
Query Match	
Query No. 22.2%; Score 502; DB 2; Length 1743;	
Best Local Similarity 27.6%; Pred. No. 2,2e-24;	
Matches 162; Conservative 78; Mismatches 158; Indels 190; Gaps 24;	
QY	19 LYGGGVAVVADPTLDR---NDS-----SMHEAFREKITALVALGTTNANPNNGH 68
DB	314 LRGGQGLTSLADTGLDSHCFFSDSKYPIPLNSVNLNR-KVVTYTTSTSDSDKVDGH 372
QY	69 GTHVAGSVLG---NALN---KGMAPQNLVFSQISWDSGGLGSLPSNINTLFSQAMNA 120
DB	373 GTHICGSAAGTPEDSVNISSFGLATDAKIAFPDLASGSSSLTP-PSDLKOLYPLXDA 431
QY	121 GARHTNSKGA---PVGATYANSRCQVDYV-RANDTVLFPAGNENGPNSGITS--APG 173
DB	432 GARVHCMSGWSVVEGYTGSISDPTASIDPLFTHPDFTILRAAGN---NEOYLSTLTQS 488
QY	174 TAKNAITVGA---NYRSPFG-----
DB	489 TAKVITVGAQITHEMYVLTDPVINYQSSVDINQSLICDPDSRYCNVTTAOCCELSNA 548
QY	193 -----SIND-----NPVHIAQFSSRGATRDGRITPDVTAQPTFLISA 229
DB	549 TTGLASCCPTLRLKRSVIDANTOPLTYNENNICSFFSKGPTHDGRMPALVARGEYITSA 608
QY	230 RSSLA-----PDSSPMANYSKYAVYGTSMATPIVAGVVAOLREH-----FIK 273
DB	609 RSNGANITDQCGDSL-PNNMALIA-ISTGSMATSPAAATTLIKOYLVUGVYPTGSIVE 666
QY	274 NRGITPKESLIKALVIGA-----TDVGLGYPSGD-----QGMGRVTLTD 312
DB	667 SNKQPTGSLIKALMINNAQLNGTFLITSSITVPSNQVFNFAAGSLVQGMGALRMS 726
QY	313 KSLVAVVYNEAT-----ALTTGQKATYSF----- 336
DB	727 NMLHVANNNSNNNKTSDGITKFDGIGGLDLVLPKNQKESLSLTGQVTSICFTYKPS 786

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G83753
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1,799 <STO>
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04550.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: vpr
C:Superfamily: microbial serine proteinase vpr; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:1-29/Domain: signal sequence #status predicted <SIG>

Query Match 13.5%; Score 304.5; DB 2; Length 799;
Best Local Similarity 23.9%; Pred. No. 4.9e-12;
Matches 145; Conservative 59; Mismatches 152; Indels 263; Gaps 28;

QY 18 GLVGGGVAVADTGLDTGRNDSMEAFRGKITALVALGR-----TNN-----AN 63
DB 171 GYTGEGITVALDITGVDTYHPD-----LVHAFGDYKQWPFIDNDDPQETPRG 218
QY 64 DPNQ---HGTHVAGSVLGNALNKMAPQANLVFQSIIMSSGGLGLPSNLT--FSQA 117
DB 219 DPREITTHGTHVAGVTAAGVLAAGLKGAPDANLLAYRL---GPGRGSTAGVAGIERA 274
QY 118 WNAGARHTNSWGPVAVGATYANRQVEYVRNNDMTVLFAAGNEPNSGTSIAPGTAKN 177
DB 275 VQDADIMNLSLGTNLNDPFAISALD--WAMEGVAVTNSNGSPNMTVSGPSTSD 333
QY 178 AITYGAT-----ENYRPSFGST----- 194
DB 334 AIVGATRLPYNKYKASVFTSDGIDYPSADIMGFPSDELLDLGETYEYAFAGLKPDP 393
QY 195 -----ADNP----- 198
DB 394 FEGDVSGKIALIVRGHIFPEYKAKENAKAGAVGAIYNNVAVGQPTVPGIAFTIMLSN 453
QY 199 -----NHIAQPSRG--ATRDRIKPDVATPGLTISA 229
DB 454 EDGLKNELENGONTVTFSLIEPKLVGETVADFPSSRGMHTMMLKPDVASGVAIVST 513
QY 230 RSLAPSSFPANYSKAVWGTSNATPIVAGVADLREHFIKNGITPKSLIKALIL 289
DB 514 IPTHQPDPI--GYGSRQ---GTSNASEHVAQALALLEAH--PNNGV---DHVRAALM 562
QY 290 AGATDV---GLGYPSGDQGWGRVTLDKSLNVAVNEATALTG----- 329
DB 563 NTAENLVDENGNRYPNHTQAGSIRI-----VDALIESETLVTPGSHSFGFTKRGQVE 617
QY 330 -----OKATYFQTO--AGKP-----LKISLVWTDAPSGTTSYTLVNDLVIRA 373
DB 618 RQHTTNLSNKRKITYPDVQFAGNPDGIKYSKILRVPGKT----- 661
QY 374 PNGQRYVN-----DFSYPYDNN--WDGNNV--NVEINA---PQSGTYIEV-- 415
DB 662 ---QKINFNVQVADARKLDPGX--VEGTIVSDGSQIVVPTILVSESDYRVTTFDLD 717
QY 416 -----QAYNVSPGPRFSLAI 431
DB 718 ENGVLFGSAY--LPGAAEFGLMI 739

RESULT 7
T28159
Pyrolysin (EC 3.4.-.-) - *Pyrococcus furiosus*
C:Species: *Pyrococcus furiosus*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000
C:Accession: T28159
R:Voorhorst, W.G.B.; Eggen, R.T.L.; Geerling, A.C.M.; Platteuw, C.; Sizeren, R.J.; Vos,
J. Biol. Chem. 271, 20426-20431, 1996

A:Title: Isolation and characterization of the hyperthermostable serine protease, pyroly-
A:Reference number: Z20481; MUID:56355370; PMID:8702780
A:Accession: T28159
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1,1398 <VOO>
A:Cross-references: EMBL:U55835; NID:g1556462; PID:g1556463; PIDN:AAB09761.1
A:Experimental source: DSM3638
C:Genetics:
A:Gene: pls
C:Keywords: hydrolase; serine proteinase

Query Match 13.4%; Score 303; DB 2; Length 1398;
Best Local Similarity 28.1%; Pred. No. 1.3e-11;
Matches 128; Conservative 37; Mismatches 154; Indels 136; Gaps 15;

QY 21 GQGVAVVADTGLDTGRNDS-----SMHEFRGKITALVALGRTNANPN-- 66
DB 301 GNGDIALVDITDIDTDEYPLQVNVTVYDAVFSTYVGLN--YLAEL---DENG 354
QY 67 -----GHGTHVAGSVLGNALN-----KG 84
DB 355 YAVFGMDGHGTHVAGVAVGVDNNDAMDMLSVSGEMVEFRLYGMDYTNVTTDTVQ 414
QY 85 MAPQANLVFQSIIMSSGGLGLPSNLTLSQANAGARHTNSWG--AVNGAYTANR 142
DB 415 VAPGAQIALRVLASDG--RGSWMDIEGMYATRHADVISMSLGNADYLDGTPESV 472
QY 143 QVDEVVRNNDMTVFAAGNEPNSGTSIAPGTAKNATVGAATE----- 185
DB 473 AVDELTEKYGAVFIAAGNEPNSGTSIAPGTAKNATVGAATE----- 185
QY 186 ---NYRPSFGSINDNPNHIAQPSRGATRDGRIKPDVTPGFIISARSLAPDSFWA 241
DB 533 YVGFYPPAYTNV-----RIAFPSRGRPRIDGELKPVVAPGVGIYSLPMWIGADF-- 585
QY 242 NYNSKYAAMGTSNATPIVAGVADLREHFIKNGITPKSLIKALILAGT-----DV 295
DB 586 -----MSGTSNATHTVSGVAVALLSG--PKPEGIYNNDDIKVLESGATWEGDPYT 636
QY 296 GLGYPSGDQGWGRVTLDKSLNVAVNEATALTGQKATYFQTOAGKPLKISLVWTDAPG 355
DB 637 GQKTELDQGGVAVVWVKSMEI-----LKAINGTTLPIVHWADKSY 678
QY 356 STASVTLVNDLVDLITAPNG-----QRYVGN 382
DB 679 SDFAETLGV--DVIRGLIARNSIPDIVEMHKKVGD 712

RESULT 8
S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - *Xanthomonas campestris* pv. c
N:Alternate names: subtilisin-related proteinase
C:Species: *Xanthomonas campestris* pv. *campestris*
C>Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 03-Dec-1999
C:Accession: S11890
R:Liliu, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A:Title: A multipurpose broad host range cloning vector and its use to characterise an
A:Reference number: S11890; MUID:90251253; PMID:2187155
A:Accession: S11890
A:Molecule type: DNA
A:Residues: 1,580 <LIU>
A:Cross-references: EMBL:X51635; NID:948533; PIDN:CAA35962.1; PID:948534
A:Experimental source: *Xanthomonas campestris* pv. *campestris*
A>Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-A1.
C:Superfamily: subtilisin; subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-33/Domain: signal sequence #status predicted <SIG>
F:168-423/Domain: subtilisin homology <SHT>

Query Match 13.1%; Score 296.5; DB 2; Length 580;
Best Local Similarity 29.0%; Pred. No. 1e-11;

Matches 137; Conservative 52; Mismatches 148; Indels 135; Gaps 26;

QY 21 GGGGVAAVADTGL-----DTGRDSSMHEAFRGKITALYALGRTNNAND-----64

DB 168 GSGTVAIVADTGTISHADLANIILAGYDPISDATTARPDNGRDSNAADGCVYAAINECGA 227

QY 65 -----PNGGTHVAGS-----VYGNALINKGMAPQANL-----VFQSIMDS 99

DB 228 GTPAASSMAGTHVACTVAIVNTNTGAGTAYGAKVYVRYLGLCGGSLIPALAIYWA 287

QY 100 SGG-IGGLPSNIN--TLFSQAMNAGARIHTNSMGAPVNGAYTANSROYDEVYRNNDMTVL 156

DB 288 SGGTVAIGIPANAPAVINNSLGGGSCSTWQNA--INGAVSRGT-----TVV 334

QY 157 PAAGNGEPN--SGTISAPGAKAIVTGATENRPSFGSADNPNIHAQSSRGATRDRI 215

DB 335 VAAANDASVSG--SLPANCANVIAVAIT--TSAGA-----KASISNFGT-----375

QY 216 KPDVTAPGTFILSARSS--LAPDSSFWANYSKYAYMGSTMATPIVAGNVAAQLREHPIK 273

DB 376 GIDVSAFGSILSTLNSGTTTSGASASYN-----GTSMASSPHVAGVVALVQS--VA 426

QY 274 NROITK--PSLTK--AALIAAGTDVGLG-----PSGQGGRTVLDKS 314

DB 427 FYALTPAVETTLKNTARALPGACSGCGAGIVNADAVTAALINGSGGGGGGNTLNG 486

QY 315 LNVAYVNEATALTGGKATYSPOTAGKPLKSLVWTDAPGTTASTYLNDLVI---371

DB 487 TPTVGLGA---TGAEINATITIVPAGSG--TLVTTTSGGS-----GADLYVRAG 531

QY 372 TANQGVKIVNDSPYDDNNWGRNVENVFIAPSGSYTTEVQAYNPSPG 423

DB 532 SAFTDSAYT-----CRPYRS-----GNAETCTITAP--SGTYRKLKAYSTPSG 573

RESULT 9

T29090
surface layer-associated STABLE proteinase - Staphylothermus marinus
N/Alternate names: hyperthermostable proteinase
C/Species: Staphylothermus marinus
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #ext_change 02-Sep-2000
C/Accession: T29090
R/Mayr, J.; Lipas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.
Curr. Biol. 6, 739-749, 1996
A/Title: A hyperthermostable protease of the subtilisin family bound to the surface layer
A/Reference number: Z20559; MUID:96385442; PMID:8793300
A/Accession: T29090
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1345 <MAY>
A/Cross-references: EMBL:U57968; NID:g1374755; PID:g1374756; PIDN:AA802323.1
A/Experimental source: strain F1
C/Function:
A/Description: probably serves an exodigestive function related to the organism's energy
A/Note: stoichiometric S-layer component

Query Match 12.7%; Score 288; DB 2; Length 1345;
Best Local Similarity 25.5%; Pred. No. 1.2e-10;
Matches 120; Conservative 66; Mismatches 152; Indels 130; Gaps 20;

QY 46 FRGKITLVALGRTNNANDPNGHGTVA-----GSVLGNALN-----KGMAPQANLV 92

DB 445 YQGRVYLA-----VSDFHGHTSVATVYASRGVLYDLGDKLYRIMGVAPGAKI- 495

QY 93 FQSIMDSGGGLGGLPSUNTLTFSQAMNAG-----ARHTNSW 129

DB 496 -----AGDPAWLLGNI--LVLEPAWLAGENITVEEDGVYVSLDPFGHRADITISNS 546

QY 130 GA-----PVNGAYTANSROYDEVYRNNDMTVLPAAGNBPNSGTISAPGTAK 176

DB 547 GSIVINFWLQGFPGIDYRSSFMDIILAIRNYLIGDHVTIVFAAGNEGPYSSNCAFGTGL 606

QY 177 NAIIVGATE--NYPSPFGSIAIDNPNIHAQSSRGATRDRIKPDVTAPGTFILSARSSLA 234

DB 607 LVITAGASTIMDTTRYLYGEGYADAVIPISSSGPFGQKPRDYNIGAF-----657

QY 235 PDSSFWAN-----YNSKYAYMGSTMATPIVAGNVAAQLREHPIKNTGITKPSLIYA 286

DB 658 -----ENASTRTIDRGYGAQPDVFGGTSEATPTSSGLALVPOAYKEVNTTDDPYTAKI 713

QY 287 ALIAGATDVGLGYPGQGWGRVTLDSLVAVYVNEATA-LTGGOKATY--SFQOAGKP 343

DB 714 ILKSAADI--WTPAPGSGGRVDALKADPTVISEWLAIVSGIOEAFLENTYTDGFPY 771

QY 344 LKISLVN--TD-----APGSTASYTL--VNDLIVTAPNG--QKYVNDSPYPDN 390

DB 772 IGYVLPYADTDYGVVYKPGS--SKNFTLNVGGAVALSMNTVLYKEYTVYDGVYDYG 830

QY 391 -----NMDGRNNT-----NVRINAPQSGTYTIEVQAYN 419

DB 831 LFLKVPKRAYSGADYEVVYVQLEBNMTPPGVFLKTPIDPLHAILISAYD 880

RESULT 10

JC4908
alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
N/Alternate names: subtilase
C/Species: Alteromonas sp.
C/Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #ext_change 08-Oct-1999
C/Accession: JC4908
R/Tsujibo, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y.
Biosci. Biotechnol. Biochem. 60, 1284-1288, 1996
A/Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacterium *Alteromonas*
A/Reference number: JC4908; MUID:97141200; PMID:8987544
A/Accession: JC4908
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-715 <TSU>
A/Cross-references: DBJ:D36600; NID:g1536787; PIDN:BA18912.1; PID:d1019647; PID:g2160
A/Experimental source: strain O-7
C/Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sensitized subtilisin.
C/Genetics:
A/Gene: aprI
C/Superfamily: subtilisin homology
C/Keywords: hydrolase
F.1.40/Domain: signal sequence #status predicted <SIG>
F.1.40/Domain: amino-terminal propeptide #status predicted <AMP>
F.151-456/Product: alkaline serine protease I #status predicted <MAT>
F.151-456/Domain: subtilisin homology <SST>
F.1497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F.239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 11.9%; Score 269.5; DB 2; Length 715;
Best Local Similarity 25.6%; Pred. No. 7.8e-10;
Matches 125; Conservative 44; Mismatches 164; Indels 155; Gaps 20;

QY 21 GGGGVAAVADTGL-----LDTRDSSMHEAFRGKITALYALGRTNNAND-----65

DB 182 GGGVVAAVLDYGRPHDLDLNIIIPGYDMISNTFVANDGARGDADDPGDAVTRGCGT 241

QY 66 -----NGGTHVAGSVLGNALN-----KGMAPQANLVFOSIMSGGLGLP 107

DB 242 DSSGQPVPRADQDSKMGTHVAGTVAIVTNGEVAIVADAKVYPRVY---GKCGGLT 298

QY 108 SNLNTLFSQAMNAGARIHTNSMGAPV-----NGAYTANSROYDEVYRNNDMTVLPAAG 160

DB 299 SDIADGIITWASGSDRPANAPVAVVIMSLGGGAGASATTQNAINQARNNGTIVIAAG 358

QY 161 NEGPNSGTISAPGAKAIVTGATENRPSFGSADNPNIHAQSSRGATRDGR-----214

DB 359 NDNDNSA-----NYPSPG-----NONGVYNAVSVG--RDGSRATYSSN 392

QY 215 --IKPDVTAPGTFILSARSSLAIP-----SSFANYSKYAYMGSTMATPIYAGN 263

DB 393 YGANIDVAPG-----GASFPADDPGILSTHNSGASPSNDSYHYEGGTSMAPHAVAGV 447

QY	264	VAGLRHEHIXKRGITPKRSLIKALLIAGATDVGIGPBGDDGMRVTLDKSLNAAVNEA	323
Dd	448	AALIKQ-----AKPSATPEVEETILKNKTTSFAGSCSNCSTGVADAA--AAVNEA	495
QY	324	--TALT--TGOKATYSEFOTQAGKEPLKISLIVTDAPGSTASYT	362
Dd	496	LGDVVTPPTGTLEDGVAKTGLSGAAGSNQFFTFDPVPAKK-----TNVTFT	541
QY	363	L--VNDDLVIITAPNGQKYVGNDFSYPDNNMDGR---NNVENFLINAPQSSTYTTEV	415
Dd	542	MSGSTGDADLIYVK-----LG---SQPTSSSYDCRPYEGGAENAECSFDAPAQTGYHMI	591
QY	416	QAAYNPFG	423
Dd	592	NGRYAYSG	599

RESULT 11

high-alkaline serine proteinase (EC 3.4.21.1), precursor - *Bacillus alcalophilus* (strain N3)
N/Alternate names: subtilisin homolog, high-alkaline
C/Species: *Bacillus alcalophilus*
C/Date: 12-Mar-1994 #sequence_revision 24-Feb-1995 #text_change 20-Jun-2000
C/Accession: A49778; JCI1244
R/Inv. Environ. Laan, J.C.; Gerritse, G.; Mulleners, L.J.S.M.; van der Hoek, R.A.C.; Quax, W.J.
A>Title: Cloning, characterization, and multiple chromosomal integration of a *Bacillus* a
A/Reference number: A49778; MUID:91282483; PMID:2059048
A/Accession: A49778
A/Molecule type: DNA
A/Residues: 1-380 <VAN>
A/Cross-references: GB:ME5086; NID:g142456; PIDN:AAA22212.1; PID:g142457
A/Experimental source: strain P392, ATCC 31408
A/Note: amino end of mature protein confirmed by peptide sequencing
R/Itakami, H.; Kobayashi, T.; Kobayashi, M.; Yamamoto, M.; Nakamura, S.; Aono, R.; Horikawa
R/Keywords: Biochem. 56, 1455-1460, 1992
A/Title: Molecular cloning, nucleotide sequence, and expression of the structural gene for
A/Reference number: JCI1244; MUID:93043753; PMID:1368952
A/Accession: JCI1244
A/Molecule type: DNA
A/Residues: 1-195, 'S', 197-380 <TK>
A/Cross-references: GB:D1315; NID:g216231; PIDN:BA02442.1; PID:g216232
A/Experimental source: *Bacillus alcalophilus* Veder, ATCC 21522 (*Bacillus* sp. 221)
C/Suprafamily: subtilisin, subtilisin homolog
C/Keywords: hydrolase; serine proteinase; zymogen
F.1-27/Domain: signal sequence #status predicted <SIG>
F.28-111/Domain: activation peptide #status predicted <PRO>
F.112-380/Product: alkaline serine proteinase #status predicted <MAT>
F.1134-340/Domain: subtilisin homolog <SBT>
F.1143,173,326/Active site: Asp, His, Ser #status predicted

```

Query Match      11.9% Score 269; DB 2; Length 380;
Best Local Similarity 31.8%; Pred No 3.6e-10;
Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;

QY AONNYGLYGOGVVAADTGLDTGRNDSMEAFEGKLTALYALGRTNANDPNHGCHTV 72
Db 126 AAHNGLTGSGYKVAVLDGTIST-----HEDLNIRGASGVPEEP-STDGNHGCHTV 177
QY AGSV--LGNALN-KMAPQANLVPOSIMDSGGTGLGSNTLTLPQAMNAGARIHTNSW 129
Db 178 AGTIALNNISIGVLGVAPPAELVYAKVLGASG--SGVSLSIQGLEMGANNGMRYANSL 235
QY GAPVAGATANGROYDEVYRNNDMTVLFPAAGEGNSGRTSPGAKAALTGATENRP 189
Db 236 GSP---SPSALTEQAIVNSKTSRGVTLVAASNGS--AGSIYPAYANMAVGAL----- 285
QY SFGSTADNDENPIAQSSRCGATRDGRIKPDVTNPGFTIILARSGLPADDSFWANNISKYAT 249
Db 286 -----DONNNAASPSQYAGL--DIVAPGVNVOSTVP-----GSTAS 321
QY MGSTSMATPIVAGNVA-----QLREHFINKNRGITPEPSELKAALLIAGATDVGL 297

```

Db	322	LNQTSNATHTVAGAAALVKQKNPSMSNVQIRNH-LKN-----TATSLGSLNL--	367
QY	298	GYPSG 302	
Db	368	-YGSQ 371	

RESULT 12

subunitisin (EC 3.4.21.62) 1 precursor - *Bacillus* sp. (strain TA39)
 C.Species: *Bacillus* sp.
 C.Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Jun-1999
 C.Accession: S23407
 R.Narinx, E.; Davail, S.; Feller, G.; Gerday, C.
 Biochim. Biophys. Acta 1131, 111-113, 1992
 A.Title: Nucleotide and derived amino acid sequence of the subunitisin from the antarctic
 A.Reference number: S23407; MUID:92256481; PMID:1581352
 A.Accession: S23407
 A.Molecule type: DNA
 A.Residues: 1-420 <NAB>
 A.Cross-references: EMBL:X62369; NID:G40200; PIDN:CAA44227.1; PID:G40201
 C.Genetics:
 A:Gene: sub1
 C:Superfamily: subunitisin; protein; subunitisin homology
 C:Keywords: extracellular protein; hydrolase; serine proteinase
 F.1-23/Domain: signal sequence #status predicted <SIG>
 F.24-111/Domain: propeptide #status predicted <PRO>
 F.112-420/Product: microbial serine proteinase #status predicted <MAT>
 F.136-374/Domain: subunitisin homology <SBT>
 F.145,185,350/Active site: Asp, His, Ser #status predicted

Query Match 11.7%; Score 265.5; DB 1; Length 420;

Matches 92; Conservative 46; Mismatches 120; Indels 59; Gaps 15;

OV 21 GOGOVAVADTGLDTGRNDSSMHEAFRCKITALL--YALGRT---NNANDPNGHGTIVAGS 75

Db 136 GCGTNTAVI DTCAVNTN - - - HDPI.PNNVECKRDEMTVGTVTNNSCTDPOGHGTHVAGS 189

CITY OF NEW YORK
OFFICE OF THE COMPTROLLER

RECEIVED MAY 19 1978

NYC- 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1. **Introduction**

2. **Background**

3. **Method**

4. **Results**

5. **Conclusion**

6. **References**

7. **Appendix**

8. **Table 1**

9. **Table 2**

10. **Table 3**

11. **Table 4**

12. **Table 5**

13. **Table 6**

14. **Table 7**

15. **Table 8**

16. **Table 9**

17. **Table 10**

18. **Table 11**

19. **Table 12**

20. **Table 13**

21. **Table 14**

22. **Table 15**

23. **Table 16**

24. **Table 17**

25. **Table 18**

26. **Table 19**

27. **Table 20**

28. **Table 21**

29. **Table 22**

30. **Table 23**

31. **Table 24**

32. **Table 25**

33. **Table 26**

34. **Table 27**

35. **Table 28**

36. **Table 29**

37. **Table 30**

38. **Table 31**

39. **Table 32**

40. **Table 33**

41. **Table 34**

42. **Table 35**

43. **Table 36**

44. **Table 37**

45. **Table 38**

46. **Table 39**

47. **Table 40**

48. **Table 41**

49. **Table 42**

50. **Table 43**

51. **Table 44**

52. **Table 45**

53. **Table 46**

54. **Table 47**

55. **Table 48**

56. **Table 49**

57. **Table 50**

58. **Table 51**

59. **Table 52**

60. **Table 53**

61. **Table 54**

62. **Table 55**

63. **Table 56**

64. **Table 57**

65. **Table 58**

66. **Table 59**

67. **Table 60**

68. **Table 61**

69. **Table 62**

70. **Table 63**

71. **Table 64**

72. **Table 65**

73. **Table 66**

74. **Table 67**

75. **Table 68**

76. **Table 69**

77. **Table 70**

78. **Table 71**

79. **Table 72**

80. **Table 73**

81. **Table 74**

82. **Table 75**

83. **Table 76**

84. **Table 77**

85. **Table 78**

86. **Table 79**

87. **Table 80**

88. **Table 81**

89. **Table 82**

90. **Table 83**

91. **Table 84**

92. **Table 85**

93. **Table 86**

94. **Table 87**

95. **Table 88**

96. **Table 89**

97. **Table 90**

98. **Table 91**

99. **Table 92**

100. **Table 93**

101. **Table 94**

102. **Table 95**

103. **Table 96**

104. **Table 97**

105. **Table 98**

106. **Table 99**

107. **Table 100**

108. **Table 101**

109. **Table 102**

110. **Table 103**

111. **Table 104**

112. **Table 105**

113. **Table 106**

114. **Table 107**

115. **Table 108**

116. **Table 109**

117. **Table 110**

118. **Table 111**

119. **Table 112**

120. **Table 113**

121. **Table 114**

122. **Table 115**

123. **Table 116**

124. **Table 117**

125. **Table 118**

126. **Table 119**

127. **Table 120**

128. **Table 121**

129. **Table 122**

130. **Table 123**

131. **Table 124**

132. **Table 125**

133. **Table 126**

134. **Table 127**

135. **Table 128**

136. **Table 129**

137. **Table 130**

138. **Table 131**

139. **Table 132**

140. **Table 133**

141. **Table 134**

142. **Table 135**

143. **Table 136**

144. **Table 137**

145. **Table 138**

146. **Table 139**

147. **Table 140**

148. **Table 141**

149. **Table 142**

150. **Table 143**

151. **Table 144**

152. **Table 145**

153. **Table 146**

154. **Table 147**

155. **Table 148**

156. **Table 149**

157. **Table 150**

158. **Table 151**

159. **Table 152**

160. **Table 153**

161. **Table 154**

162. **Table 155**

163. **Table 156**

164. **Table 157**

165. **Table 158**

166. **Table 159**

167. **Table 160**

168. **Table 161**

169. **Table 162**

170. **Table 163**

171. **Table 164**

172. **Table 165**

173. **Table 166**

174. **Table 167**

175. **Table 168**

176. **Table 169**

177. **Table 170**

178. **Table 171**

179. **Table 172**

180. **Table 173**

181. **Table 174**

182. **Table 175**

183. **Table 176**

184. **Table 177**

185. **Table 178**

186. **Table 179**

187. **Table 180**

188. **Table 181**

189. **Table 182**

190. **Table 183**

191. **Table 184**

192. **Table 185**

193. **Table 186**

194. **Table 187**

195. **Table 188**

196. **Table 189**

197. **Table 190**

198. **Table 191**

199. **Table 192**

200. **Table 193**

201. **Table 194**

202. **Table 195**

203. **Table 196**

204. **Table 197**

205. **Table 198**

206. **Table 199**

207. **Table 200**

208. **Table 201**

209. **Table 202**

210. **Table 203**

211. **Table 204**

212. **Table 205**

213. **Table 206**

214. **Table 207**

215. **Table 208**

216. **Table 209**

217. **Table 210**

218. **Table 211**

219. **Table 212**

220. **Table 213**

221. **Table 214**

222. **Table 215**

223. **Table 216**

224. **Table 217**

225. **Table 218**

226. **Table 219**

227. **Table 220**

228. **Table 221**

229. <

128 S W A P V N G A I I A N S R Q V D E I V K N N D M I V L F A H G N E G F N S G I I S A F G I A R V A I I V A R L E N I 10

Db 249 SLGSSGESSLITNA---VNYSYNGVLI AAGNSGPYQGSIGYPGALVNNAVVALEN- 304

188 RPSFGSIADNPNHIAQFSSRG-----ATRDGRIKPDVTAPGTFILSARSSLAPDSSF 233

Db 305 -----KVENGTYRVADFSSRGYSWTDGDYAIQKGV--EISAPGAITYST----- 347

QY 240 WANNYSKAYMGTSMATPIVAGNVAQLREHFIKNRGITPKPSLJKAALIAGATDVGLGY 295

348 W--FDGGYATISGTSMA SPHAGLAKIWA OYPSASNVDVRGELOYRAY---ENDILSGY 402

200 PSC-111 DOCUMENT 311

..

RESULT 13

US0173
alkaline proteinase (EC 3.4.21.-) A precursor - *Vibrio alginolyticus*

N;Alternate names: alkaline serine exoproteinase A
C;Species: *Vibrio alginolyticus*

C:\Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

R;Deane, S.M.; Robb, F.T.; Robb, S.M.; Woods, D.R.

A;Title: Nucleotide sequence of the *Vibrio alginolyticus* calcium-dependent, d

A:Accession: JS0173
A:Reference number: JS0173; MOLD:89326126; PMID:2346861

A;Molecule type: DNA

A:Residues: 1-534 <DEA>
 A:Cross-references: GB:M25499; NID:G155250; PIDN:AA27550.1; PID:G155251
 C:Comment: This protein is a calcium-dependent and sodium dodecyl sulfate-resistant pro-
 C:Superfamily: subtilisin; subtilisin homology
 C:Keywords: calcium; hydrolase; serine protease
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:1-22-534/Product: alkaline serine proteinase A #status predicted <MAT>
 F:171-378/Domain: subtilisin homology <SBT>
 F:1271-273,297-299/Region: S1 specificity crevice #status predicted
 F:180,213,363/Active site: Asp, His, Ser #status predicted

Query Match 11.7%; Score 265; DB 1; Length 534;
 Best Local Similarity 26.3%; Pred. No. 1e-09;

Matches 123; Conservative 65; Mismatches 152; Indels 128; Gaps 24;

```

QY 7 IYKADVAQ-----NNY--GLYGGQGVAVADGLDGRDSSMHEAF 46
DB 137 IYADANQNALWGLDRIDQRLPLDNNYSANFDGVTAYIDTV-----NNAHYEF 190
QY 47 RGRITALLY-ALGRTNNANDPNHGHTVAGSVLGNALNKGMAPANLVFQSIM--DSGGL 103
DB 191 GGRSVGVDPEVDADADPCNGHGTVAHTIGSLY--GVAKNVLVGVRLVSCGSGST 248
QY 104 GGLPSNLNTLFSQANAGARIHTNSGCAVNGAYTANSRQVDEYVNNMTVLPAAGNG 163
DB 249 SGVIAGDVWVAANA--SGSVANMSLG--GGQVALDSAVQAVQSG-VSFMLAAGNSN 302
QY 164 PNEGTSAPGTACATVGTATENYRPSFGSIADPNHIAQFSRGATRDGRITKPYTAPG 223
DB 303 ADACNYS-PARVATGVTVGST-----TSTDARSSFSWMSGSC-----VDVFAPG 344
QY 224 TFLLSARSSLAPDSSEFWANYSKAYATMGTSNATPIYAAGVAQLRHFIRKNGITPKESL 283
DB 345 SQIKSA-----W-YDGGYKXTSGTSMATPVAG-VVAL-VLDENSSVS--PSQ 387
QY 284 IKALILAGA-----TDVGLGYPSG-----DQGMGVTLDKSLINVA 319
DB 388 VEHIVRASTGKVTDRGSVNLILSLTDADCGQCGGDPPTPEDEGLTSGVPVS--- 444
QY 320 VNEATALTGGQKATYSFQIQAGKPLKISLVMTDAPSGTASLYLVNDLVLITAPNGQKY 379
DB 445 ---GLSGSSQGVAYVYDVAGQRLTVQM-----YGGSDADLVYRFR--GAK- 486
QY 380 VGNDFSPYDNMWDGR---NNVENVFINAPSGGTITIEVOAYNVEPSG 423
DB 487 -----PTINAMDCRPFKYGNETCVSKATQSGRIHVMIGYSNMSG 527

```

RESULT 14

S25835
 Subtilisin (EC 3.4.21.62) precursor - *Bacillus* sp. (strain TA41)
 C:Species: *Bacillus* sp.
 C:Date: 22-Nov-1993 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999
 C:Accession: S25835
 R:Davidall, S.; Feller, G.; Narinx, E.; Gerday, C.
 Gene 119, 143-144, 1992
 A:Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph *Bacillus*
 A:Reference number: S25835; MUID:93012966; PMID:1398082
 A:Accession: S25835
 A:Molecule type: DNA
 A:Residues: 1-419 <DAV>
 A:Cross-references: EMBL:X65533; NID:G40198; PIDN:CAA5096.1; PID:G40199
 C:Superfamily: subtilisin; subtilisin homology
 C:Keywords: extracellular protease; hydrolase; serine protease
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-110/Domain: signal sequence #status predicted <PRO>
 F:111-419/Product: microbial serine proteinase #status predicted <MAT>
 F:115-373/Domain: subtilisin homology <SBT>
 F:144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 11.4%; Score 258.5; DB 1; Length 419;
 Best Local Similarity 31.1%; Pred. No. 2e-09;
 Matches 82; Conservative 40; Mismatches 99; Indels 43; Gaps 12;

```

QY 21 GGGVVAVADTGLDGTGRDSSMHEAFRGRITALLYALGR---NNANDPNHGHTVAGSV 76
DB 135 GAGINIAVDTGVYTNHNPDLSSNVEQCKD-----FVGTNFTDNCSTDRQGHGTVAAGSA 189
QY 77 LGNALN---KMAPANLVFQSIM--DSGGIGLPSNLNTLFSQANAGARIHTN--SW 129
DB 190 LANGTGSGVYGVAPADLWAVYVLGDGSGVADIDAEALRHAGDQATALTNTKVTINMSL 249
QY 130 GAPVNGAYTANSRQVDEYVNNMTVLPAAGNPNSTGISAPGTAKNATVGTATENYRP 189
DB 250 GSGGESSLITNA--VD-YADKGVLLIIAAGNCPKPSGTYGALVNAVVAALN--- 303
QY 190 SFGSIADPNHIAQFSRGATRDG-----RIKPYTAPGTFLLSARSSLAPDSFWAN 243
DB 304 ---TIQGTAVADFSSRGHKTAGDYVIQKGDVEISAPQAAVYST-----W--F 348
QY 244 NSKAYVNGTSMATPIYAAGVAQL 267
DB 349 DGGYATISGTSMSAPHAAGLAAKI 372

```

RESULT 15

D75393
 serine protease, subtilase family - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: D75393

R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathavan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Mc
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: D75393

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-627 <MHI>

A:Cross-references: GB:AE001990; GB:AE00513; NID:G6459214; PIDN:AAFL1026.1; PID:G64592;

A:Experimental source: strain R1

A:Gene: DR1459

A:Map position: 1

Query Match 11.3%; Score 256.5; DB 2; Length 627;

Best Local Similarity 24.8%; Pred. No. 4.5e-09;
 Matches 104; Conservative 61; Mismatches 161; Indels 93; Gaps 17;

```

QY 18 GLYGGQGVAVADTGLDGTGRDSSMHEAFRGRITALLYALGR---NNANDPNHGHTV 72
DB 161 GFKGQNTIRIHLDBGIDPS-----HPELAGKVAAPGEPNGEGDRVSSQPHDTTGHHT 214
QY 73 AGSVLGNALNKGMAPANLVFQSIMDSSG-----LGLPSNLNTLFSQANAGARIHTN 127
DB 215 AGLVGSKV--GVAPKAVISALVLPNNEGFAVYVIMQGVLDPDNNADTDQDADVVM 272
QY 128 SWGAPVNGAYTANSRQVDEYVNNMTVLPAAGNPNSTGISAPGTAKNATVGTATENY 187
DB 273 SLGIRP--GTWNEFTVPVNNMLKAGVVPV-FALGNFGPAGSGTSPGMLPCAIGVGA--- 325
QY 188 RPSFGSIADPNHIAQFSRGATRDGRI-----KPYTAPGTFLLSARSSLAPDSFWA 241
DB 326 -----VDSNGQVASFSSRQPVAMQGEISVFTRKPIAAGVNTITVR----- 368
QY 242 NYSKYAVMGTSMATPIYAAGVAQLRHFIRKNGITPKP---SLIKALIAATDVGL 297
DB 369 --NGYQMSGSSQASPTAGAVAVL-----LSAKPASVDALIKNALFTSASASA 417
QY 298 GYPSGDQMGKVTLDKSLNAVYVNEATALTGGQKATYSFQIQAGKPLKISLVMTDAPGST 357
DB 418 --KNNNVGFGQISIFGAG-----KGVSTGTSPQPTPTTPAPPTTPTPTTPAP 469
QY 358 T-----ASYTLVNDLVLITAPNGQKYV---NDFSPYDNMWDGRNNEVNF 402

```

Sat Apr 3 17:20:03 2004

us-09-985-689a-3.rpr

Page 8

Db 470 TPPAPAKPENGPAGYTL-----CAVBSGRCSGARQKQVAFGTDGRWISGSTDDTP 520

Search completed: March 31, 2004, 16:10:17
Job time : 12.3246 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 7.49423 Seconds
(without alignments)

3008.498 Million cell updates/sec

Title: US-09-985-689A-3

Perfect score: 2263
Sequence: 1 NDVARGIVKADVAQNNNGLY.....EVQAVNPVSGRPSRLAIWH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	502	22.2	1743	1 TAGC_DICDI	Q23868 dictyostei
2	463.5	20.5	1905	1 TAGC_DICDI	P54683 dictyostei
3	332	13.8	806	1 SUBT_BACST	P29141 bacillus su
4	305	13.5	1398	1 PLS_PFRFU	P72186 pyrococcus
5	296.5	13.1	580	1 EXPR_XANCP	P23114 xanthomonas
6	269	11.9	269	1 PRTP_BACSP	Q99405 bacillus sp
7	269	11.9	269	1 SUBS_BACLE	P29600 bacillus le
8	269	11.9	380	1 ELYA_BACAO	P27693 bacillus al
9	269	11.9	380	1 ELYA_BACCS	P28162 bacillus cl
10	265.5	11.7	420	1 SUBT_BACCS	P28842 bacillus sp
11	265	11.7	534	1 PROA_VIBAL	P16588 vibrio algi
12	259	11.4	269	1 SUBB_BACLE	P29599 bacillus le
13	255	11.3	401	1 THEB_BACSP	Q45670 bacillus sp
14	248.5	11.0	1167	1 SCAL_STRPY	P15926 streptococ
15	246	10.9	894	1 WPRB_BACSV	P54423 bacillus su
16	243.5	10.8	378	1 ELYA_BACSP	P20724 bacillus sp
17	243	10.7	513	1 AQUJ_THBAQ	P08854 thermus agu
18	242	10.7	379	1 SUBT_BACLI	P00780 bacillus li
19	239.5	10.6	1181	1 SCAL_STRPY	P58099 streptococ
20	239.5	10.6	321	1 ISP_BACCS	P29140 bacillus cl
21	237.5	10.5	422	1 TKSU_PYRKO	P58502 pyrococcus
22	236.5	10.5	402	1 ALP_CEPAC	P29118 cephalospor
23	231.5	10.2	1052	1 MSIF_HUMAN	Q14703 homo sapien
24	229.5	10.1	319	1 TSP1_BACSV	P11018 bacillus su
25	228	10.1	1052	1 MSIF_CRIGR	Q92248 cricetus su
26	228	10.1	1052	1 MSIP_MOUSE	Q9W422 mus musculu
27	227.5	10.1	1052	1 MSIP_RAT	Q9W423 ratnus noxv
28	227.5	10.1	381	1 SUBN_BACNA	P35835 bacillus su
29	227.5	10.1	381	1 SUBT_BACSA	P00783 bacillus su
30	227.5	10.1	381	1 SUBT_BACST	P29142 bacillus su
31	226.5	10.0	381	1 SUBT_BACST	P29143 bacillus su
32	226.5	10.0	1433	1 SUBT_BACSV	P16397 bacillus su
33	225	9.9	382	1 SUBT_BACAM	P00782 bacillus am

34	224.5	9.9	275	1 SUBT_BACPU	P07518 bacillus pu
35	224	9.9	274	1 SUBD_BACLI	P00781 bacillus li
36	223.5	9.9	279	1 THEB_THBVU	P04072 bacteroidin
37	222	9.8	595	1 BPRX_BACNO	P42780 bacteroides
38	220	9.7	404	1 SMP1_MAGPO	Q9Y778 magnaporthe
39	219	9.7	388	1 CUDP_METAN	P29138 metathiazium
40	217.5	9.6	326	1 ISP_PABPO	P29139 paenibacill
41	213.5	9.4	409	1 ALP_TRTHA	Q03420 trichoderma
42	213	9.4	467	1 ISB6_SCHPO	P40903 schizosacch
43	210	9.3	533	1 FBP6_ASFNG	P33295 aspergillus
44	209.5	9.3	451	1 YLPI_SCHPO	Q9ut80 schizosacch
45	209	9.2	530	1 HLY_HAU17	P29143 halophilic

ALIGNMENTS

RESULT 1
TAGC_DICDI
ID TAGC_DICDI STANDARD; PRT; 1743 AA.
AC Q23868;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prestalk-specific protein tagc precursor (EC 3.4.21.-).
GN TAGC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA MEDLINE=97140317; PubMed=8986798;
RX Shaulsky G., Escalante R., Loomis W.F.;
RT "Developmental signal transduction pathways uncovered by genetic
suppressors";
RL Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).
CC -1- FUNCTION: Inter-cellular communication via tagc may mediate
integration of cellular differentiation with morphogenesis (by
similarity).
CC -1- SIMILARITY: In the N-terminal section; belongs to peptidase family
S8.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -1- SIMILARITY: STRONG, TO TAGB.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; U60086; AAB03331.1; -
DR PIR; T18279; T18279.
DR DictyBase; DDB0001795; tagc.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transport.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PS00723; SUBTILISIN.
DR SMART; SM00382; AAA_1
DR PROSITE; PS50929; ABC_TMIF; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00136; SUBTILASE_SER; FALSE_NEG.
DR Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
KM

```

KW Signal. 1 27
FT CHAIN 28 1743
FT CHAIN 316 642
FT CHAIN 1450 1687
FT TRANSMEM 962 982
FT TRANSMEM 1027 1047
FT TRANSMEM 1072 1092
FT TRANSMEM 1157 1177
FT TRANSMEM 1260 1280
FT TRANSMEM 1286 1308
FT TRANSMEM 1325 1325
FT ACT SITE 372 372
FT ACT SITE 637 637
FT NP BIND 1485 1492
FT NP BIND 42 46
FT DOMAIN 94 103
FT DOMAIN 643 646
FT DOMAIN 733 741
FT DOMAIN 786 792
FT DOMAIN 1337 1340
FT DOMAIN 1346 1352
FT DOMAIN 1353 1357
FT DOMAIN 1358 1364
FT DOMAIN 1381 1386
FT DOMAIN 1707 1729
FT CARBOHYD 390 390
FT CARBOHYD 536 536
FT CARBOHYD 547 547
FT CARBOHYD 614 614
FT CARBOHYD 689 689
FT CARBOHYD 735 735
FT CARBOHYD 741 741
FT CARBOHYD 766 766
FT CARBOHYD 832 832
FT CARBOHYD 887 887
FT CARBOHYD 1251 1251
FT CARBOHYD 1385 1385
FT CARBOHYD 1386 1386
FT CARBOHYD 1454 1454
FT CARBOHYD 1704 1704
SQ SEQUENCE 1743 AA; 194145 MW; 12DB363E2F729839 CRC64;

Query Match 22.2% Score 502; DB 1; Length 1743;
Best Local Similarity 27.6% Pred. No. 7.2e-25;
Matches 162; Conservative 78; Mismatches 158; Indels 190; Gaps 24;

19 LVGGQGVAVADTGLDGR--NDS-----SMHEAFRGKITALYALGRTNNANDPNC 68
314 LRGGQILSIADTGLDGHCHFCFSDSKYPIPLNSVNLNHR-KVVTYITTSIDSDSKVDGH 372
69 GTHVAGVLG-----NALN-----KGMAPQANLVQSIWDSGG-GLPSNINTLFSQAWNA 120
373 GTHICGSAAGTPEDSSYNISSPGLADAKIAFPDLASSSSLTLP-PSDLKQLYQPLVYA 431
121 GARLHTNSWGA---PVNGAYTANRSQVDEYV-RNNDMTVLPAAGNEGPNSTIS--APG 173
432 GARVHCDMSGVSVEGYTGYSSTASIDFLFHPDFILRAAGN---NEQYLSLDTOS 488
174 TAKAIVVGATE-----NYRPSFG----- 192
489 TAKAVITVGAHQTHENYLTLDGPNYINYGSSVDINQELCDPFSRYCNYTTACCEESNA 548
193 -----SIAD-----NPNHIAQFSSRGATRDGRIRKPDVATGTFILSA 229
549 TTGLASCCPTLRLKRSVIDAANTQPLNENNICFSSSKGPTHGMRKFPALVAGEYITSA 608
230 RSSSLA-----PDSSFANYNYSKAMGGMATPIYAGVAAQLREH-----FIK 273
609 RSNGANTTDCCGDSGL-PNTVALA-LSGTSMAISFAAAATTIRQYLVGVGYPTGSIVE 666
274 NRGITPKRSILKALLIGA-----TDVGLGVPSCD-----QSGMRVTLTD 312

```

```

Db 667 SNKQPTGSLKALMINNAQLNGTFOLITSSITPSNQVENFAGASLVQMGAIKMS 726
Qy 313 KSLNVAVNAT-----ALTTGKATYSF----- 336
Db 727 NMHVAVNNNSNNKTSQGITFSDIGLRLVYKPNQKESLSGTQWTSICFYKRS 786
Qy 337 --OTQAGKPLK--ISLVWTDAPGSTTASVTLVNDLVLITAPNGQKYGNDPSYEDN- 390
Db 787 SSSNSGNNIPRVAVTLWTDPPSYAGAKENLVNNDLTMV-----YRDNSTIFYSNQ 841
Qy 391 -----NWDGRNVENVFINAPQSGTYTTEVAYNVPGSGPQRS 428
Db 842 GGSSFFGLAPQTQTLNNVSGIVNFTPEMTYRPMVAGTVPMGPNQFS 889

RESULT 2
TAGB_DICDI
ID TAGB_DICDI STANDARD; PRT; 1905 AA.
AC P54683;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prestalk-specific protein tagb precursor (EC 3.4.21.-).
GN TAGB.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95262903; PubMed=7744252;
RX STRAIN=AX4;
RA Shaulsky G., Kuapa A., Loomis W.F.;
RT "A multistep resistance transporter/serine protease gene is required for prestalk specialization in Dictyostelium."
RL Genes Dev. 9:1111-1121(1995).
CC -1- FUNCTION: Intercellular communication via tagb may mediate integration of cellular differentiation with morphogenesis.
CC -1- SIMILARITY: In the N-terminal section, belongs to peptidase family S8.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -1- SIMILARITY: STRONG, TO TAGC.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).
CC EMBL, U20432; AAA62212.1; -.
CC PIR: T18267; T18267.
CC MEROPS: S08.006; -.
CC DictyBase; DB0001964; tagb.
CC InterPro; IPR003593; AAA_Atpase.
CC InterPro; IPR001140; ABC_TM_transp.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR002029; Peptidase S8.
CC Pfam; PF00664; ABC_membrane_1.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF00082; Peptidase S8; 1.
CC PRINTS; PR00723; SUBTILISTIN.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS50929; ABC_TM1F; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
CC PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC HydroLase; Serine protease; ATP-binding; Transport; Transmembrane; Signal.
KW Signal.

```

```

FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 1905 PRESTALK-SPECIFIC PROTEIN TAG.
FT DOMAIN 378 700 PROTEASE.
FT DOMAIN 1518 1756 ABC TRANSPORTER.
FT TRANSMEM 1011 1031 POTENTIAL.
FT TRANSMEM 1076 1096 POTENTIAL.
FT TRANSMEM 1121 1141 POTENTIAL.
FT TRANSMEM 1210 1230 POTENTIAL.
FT TRANSMEM 1309 1329 POTENTIAL.
FT TRANSMEM 1332 1352 POTENTIAL.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 432 432 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 695 695 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1553 1560 ATP (POTENTIAL).
FT DOMAIN 63 67 POLY-GLN.
FT DOMAIN 95 104 POLY-ASN.
FT DOMAIN 107 134 POLY-ASN.
FT DOMAIN 311 321 POLY-SER.
FT DOMAIN 833 837 POLY-SER.
FT DOMAIN 838 844 POLY-GLY.
FT DOMAIN 871 876 POLY-LEU.
FT DOMAIN 1012 1015 POLY-ILE.
FT DOMAIN 1386 1389 POLY-GLY.
FT DOMAIN 1398 1404 POLY-GLY.
FT DOMAIN 1445 1450 POLY-ASN.
FT DOMAIN 1765 1779 POLY-ASN.
FT DOMAIN 1782 1785 POLY-SER.
FT DOMAIN 1807 1812 POLY-PRO.
FT DOMAIN 1813 1860 POLY-GLN.
FT DOMAIN 1872 1878 POLY-PRO.
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1172 1172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1522 1522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1658 1658 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1905 AA; 212518 MW; B8E223FA8B9AE13C CRC64;

```

```

Query Match 20.5%; Score 463.5; DB 1; Length 1905;
Best Local Similarity 28.0%; Pred. No. 2,6e-22;
Matches 162; Conservative 77; Mismatches 156; Indels 183; Gaps 28;

```

```

QY 19 LVGGQGVAVADTGLDTR--NDS-----SMHEAFRGKITLALYALGRTNANDP--N 66
DB 376 LKGGQGLSLADTGLDGHCHPFSDSKYPFENQVAMENHRKVTYI-----TYHDMEDYVN 430
QY 67 GHGTHVAGSVLGN-----ALN--KMAPOALVQSLINDSGGLGL--PSNLTLFESQA 117
DB 431 GHGTHVCGSAGPDPEDSSWAISFSGLATDAXIAFYD--SSGSSEPTPPEDYSQMYKPL 488
QY 118 WNAAGAHINSMWGA-----PVNGAYTANSQVDEYV--RNNDMVLPFAAGNEGNSGTLAP 172
DB 489 YDGAARVHDSDMSVSLQGYGVGSPDADGADIAFLYEYEFSLIAAGN--NELFASLLAQ 547
QY 173 GTAKNAITVGATE---NT-----RPSF----- 191
DB 548 ATKNAITVGABGTAFHNVVSDALEYDPSDNANPQPLFPDKKYCNNTTAKCCEVSNV 607
QY 192 -----GSIND-----NPNHIAQFSRGATPRGRIRKPEVTAPGTILSAR 230
DB 608 KGIQLCCPASIKMNASDSTTQOPFENNMMGFSKGPHTHGRKLPDLVAAGEVITBAR 667
QY 231 SS-----LAPDSSEWANTNSKYAVVGTSMATPIVAGVAVOLREHF-----IKN 274
DB 668 SNGENSTDOCGDSL--PNANGILMS--ISGISMATPLATAATTLIRQYLDVGYEPTGESYEE 725
QY 275 RGTTPKPSITIKALLINGATVGLGY-----PSCD-----QGMGRVTLDSL 315
DB 726 NKLLPGLSLIKALMINNAQLINTYFWSASSTNPNSAITEQINGANLIGKALMANMWL 785
QY 316 NVAYVN-----EAT-----ALTGQKATYSPQTO-----AGK 342

```

```

DB 786 YKSSNPPTPPSMWIGIGIGKRNQKATEWKEDSLSSGLNKSICYFTYKPPSSSSGSGCGGCT 845
QY 343 P-LKISLWTDAPAGSTTASVTLVNDL-----VITARN--GQKYVGNDFSYPYDN 390
DB 846 PRLVATLVTPDPSPVSGAKFNLVNLLDLSNDDSDITITNSGSLQPAKVAQP--- 902
QY 391 NMDGRNNVENFVFNAPQSGTYTTEVQAVVPSGPPQPS 428
DB 903 --DTLNVEGIIINPTKANNYKFTIAGTVIPFGPKFS 938

RESULT 3
SUBV_BACSU
ID SUBV_BACSU STANDARD; PRT; 806 AA.
AC P29141;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Minor extracellular protease vpr precursor (EC 3.4.21.-).
GN VPR OR IPA-45R OR BSU38050.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 161-195.
RX MEDLINE=92041574; PubMed=1938892;
RA Stoma A., Rufe G.A. Jr., Theriault K.A., Dwyer M., Wilson S.W.,
RA Pero J.;
RT "Cloning and characterization of the gene for an additional
RT extracellular serine protease of Bacillus subtilis."
RJ J. Bacteriol. 173:6889-6895(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95020537; PubMed=7934828;
RA Glaeser P., Kunze F., Arnold M., Goudart M.P., Gonzales W.,
RA Hillo M.F., Ionescu M., Lubochinsky B., Marcelino L., Mosser I.,
RA Pressecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees."
RJ Mol. Microbiol. 10:371-384(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunze F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bettencourt M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Calwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.U., Connerton I.F., Cummings N.J., Daniel K.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizli A., Galleron N.,
RA Gilm S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.U., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hiltbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
RA Kunita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Melado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Oglawa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwollik S., Prescott A.M.,
RA Pressecan E., Puig P., Punnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
RA Sager T., Scallan E., Schleich S., Schreier R., Scofield F.,
RA Setiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
RA Tatenuchi M., Tanakoshi A., Tanaka T., Terstra P., Tomonori A.,
RA Torato V., Uchiyama S., Vandenberg M., Vanter F., Vassarotti A.,
RA Viati A., Wandutt R., Wedler E., Wedler H., Weitzneger T.,

```

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Darchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT subtilis".
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: NOT REQUIRED FOR GROWTH OR SPOULATION.
 CC -1- SUBCELLULAR LOCATION: Secretd.
 CC -1- PTM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; W6590; AAA2881.1; -;
 CC EMBL; X73124; CA51601.1; -;
 CC EMBL; Z99123; CAB15835.1; -;
 CC PIR; A41341; A41341.
 CC HSSP; P00782; 2SMT.
 CC MEROPS; S08; UPA.
 CC Subtilisin; BG10591; VPR.
 CC InterPro; IPR003137; PA.
 CC InterPro; IPR009020; Peptidase_S8.
 CC InterPro; IPR009020; Peptidase_inh1b.
 CC Pfam; PF02225; PA; 1.
 CC Pfam; PF00082; Peptidase_S8; 1.
 CC PRINTS; PR00723; SUBTILISIN.
 CC PROSITE; PS00136; SUBTILISIN_ASP; 1.
 CC PROSITE; PS00137; SUBTILISIN_HIS; 1.
 CC PROSITE; PS00138; SUBTILISIN_SER; 1.
 CC Hydrolyase; Serine protease; Zymogen; Signal; Complete proteome.
 KW SIGNAL 1 28 POTENTIAL.
 FT PROPEP 29 160
 FT CHAIN 161 806 MINOR EXTRACELLULAR PROTEASE VPR.
 FT ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 233 233 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 534 534 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SSOURCE 806 AA; 85608 MW; P984E3BF08B69DD CRC64;
 Query Match 13.8%; Score 312; DB 1; Length 806;
 Best Local Similarity 23.4%; Pred. No. 6, 1e-13;
 Matches 129; Conservative 56; Mismatches 156; Indels 210; Gaps 17;
 QY 18 GLYGQGVAVADTGLDGR-----NDSSMHEAFRGKITALVALGRTN 60
 DB 177 GYGKGIKVALIDTGVYVHNPDLKKNFGQYKGYDFVNDYDPEKPTG----- 224
 QY 61 NNDPNG-----HGHYAGSVLGNALNKGMAPQANLVFQSISSGGLGSPENLTLFSQ 116
 DB 225 ---DPGEATDGTGTHAGVAAAGITKGVAPDITLAVLDPGG--STTENVAGVER 279
 QY 117 AMNAGARIHTNSGAVNGAYTANSPQVDEYVYRNDMTVLPAQNEGSPGTSAPGTAK 176
 DB 280 AVQDGDADVMNLSLGNSLNPDATSPALD-WMSEGVAVVTNGNSGPGWTVGSPGRSR 338
 QY 177 NATTVATE---NYRPSFGST----- 194
 DB 339 EALISVATQPLNEVYVFGSYSSAKVWGYNKEDYKALNKEVELVEAGIGEAADFEGK 398
 QY 195 -----ADNP----- 198
 DB 399 DLTKGAVVVRGSIADVADNAKKAGAIQVYVNNLSGEIEANYPGMSVPTIKSLDGG 458
 QY 199 -----NHIAQFSKRAITP-GRKIGDVTAAPGFTLSAAS 232
 DB 459 EKVVALKAGETTKTKTLVSKALGEOVADFFSSRGVMDTWIKDISPGVNIYSTIPT 518
 QY 233 LAPDSSFWMYNSKYAVMGTSNATPIVAGNVAQLREHFIKORGITTPKSL--IKALIA 290

DB 519 HDPDHPY--GYGSKQ-----GTSMSPHIAGAVAVIKQ-----AKPKMSVEQIKALIMN 565
 QY 291 GATDV-----GLGYPSGDQGWGRVTLDKSLNAVAYNEATLTTGQKATY-SFOTACKPKPK 345
 DB 566 TAVTLKDSDEGVYPHNAQAG-----SARINMAIRKADSLVSPGSGYGFLENGENETK 619
 QY 346 ISLWTDAPSTASATLVMDLVLITAPRGQYVGNDFSPYDNNMDGRNVEYFIMA 405
 DB 620 NETFTTINOSIKRSYTL-----EYSPNGSIGISTGTSRVVIPA 658
 QY 406 POSGTYTEVQ 416
 DB 659 HQTKATATKVK 669
 RESULT 4
 PLS_PYRPU STANDARD; PRT; 1398 AA.
 ID P72186;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pyrolysin precursor (EC 3.4.21.-).
 GN PLS OR P70287.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Pyrococcus.
 CX NCBI_Taxid=2261;
 RN [1]
 RP SEQUENCE FROM N.A.A. SEQUENCE OF 150-184, AND CHARACTERIZATION.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RX MEDLINE=96355370; PubMed=8702780;
 RA Voochhorst W.G.B., Eggen R.I.L., Geerling A.C.M., Platteuw C.,
 RA Slezien R.J., de Vos W.M.;
 RT "Isolation and characterization of the hyperthermostable serine
 RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
 RT *Pyrococcus furiosus*.".
 RL J. Biol. Chem. 271:20426-20431 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.A.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the *Pyrococcus furiosus* genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP CHARACTERIZATION, AND 3D-STRUCTURE MODELING.
 RX MEDLINE=21079021; PubMed=11210516;
 RA de Vos W.M., Voochhorst W.G.B., Dijkgraaf M., Kluskens L.D.,
 RA Van der Oost J., Slezien R.J.;
 RT "Purification, characterization, and molecular modeling of pyrolysin
 RT and other extracellular thermostable serine proteases from
 RT hyperthermophilic microorganisms.";
 RL Mehl. Enzymol. 330:383-393 (2001).
 CC -1- FUNCTION: Has endopeptidase activity toward caseins, casein
 CC fragments including alpha-SI-casein and synthetic peptides.
 CC -1- SUBCELLULAR LOCATION: Cell envelope associated.
 CC -1- PTM: LHM pyrolysin seems to be produced by autoprotoeolytic
 CC activation of HWM pyrolysin.
 CC -1- PM: Glycosylated.
 CC -1- MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U55835; AAB09761.1; -;
 DR EMBL; AEO10153; AAL80411.1; -;
 DR PIR; T28159; T28159.

DR HSSP; Q45670; 1DBI.
 DR MEROPS; S08.100; -
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
 KW Complete proteome.
 FT SIGNAL 1 26
 FT PROPEP 149
 FT CHAIN 150 1398
 FT ACT_SITE 179
 FT ACT_SITE 179
 FT ACT_SITE 365
 FT ACT_SITE 590
 FT CARBOHYD 152
 FT CARBOHYD 152
 FT CARBOHYD 222
 FT CARBOHYD 228
 FT CARBOHYD 240
 FT CARBOHYD 257
 FT CARBOHYD 262
 FT CARBOHYD 298
 FT CARBOHYD 327
 FT CARBOHYD 327
 FT CARBOHYD 406
 FT CARBOHYD 406
 FT CARBOHYD 651
 FT CARBOHYD 651
 FT CARBOHYD 739
 FT CARBOHYD 739
 FT CARBOHYD 792
 FT CARBOHYD 893
 FT CARBOHYD 908
 FT CARBOHYD 917
 FT CARBOHYD 917
 FT CARBOHYD 929
 FT CARBOHYD 1048
 FT CARBOHYD 1056
 FT CARBOHYD 1084
 FT CARBOHYD 1117
 FT CARBOHYD 1117
 FT CARBOHYD 1140
 FT CARBOHYD 1140
 FT CARBOHYD 1148
 FT CARBOHYD 1208
 FT CARBOHYD 1237
 FT CARBOHYD 1237
 FT CARBOHYD 1332
 FT CARBOHYD 1332
 FT CONFLICT 607
 FT CONFLICT 881
 FT CONFLICT 881
 SQ SEQUENCE 1398 AA; 154474 MW; 355D873A27D56552 CRC64;
 Query Match 13.5%; Score 305; DB 1; Length 1398;
 Best Local Similarity 28.1%; Pred. No. 3, 5e-12;
 Matches 128; Conservative 37; Mismatches 154; Indels 136; Gaps 15;
 21 GGGQVAVADTGTGDTGRNDS-----SMHEAFRGKITATLALGRTNANDPN-- 66
 301 GNGYDLAVYDTDDYPTDEVPVLCQVNTYDVAVFYGGPLN--YVLAET----DPNCE 354
 67 -----GHGHVAGSVTGNALN-----KG 84
 355 YAVFGDGHGHVAGVYAGVYDSNNDAMDWLSMYSGEWVSRILYGMVYNTVTDVVG 414
 85 MAPQANLVFGSINDSSGGJGLPSNLTFSQAMNAGARLHTNWSG--APVNGAVTANR 142
 415 VARGAGIMAIRVRSBG--RGSMDIIEGWTVAATGADVISMISLGSNAPYLDGDPREV 472
 143 QVDEYVRANDMTVLFAAGNENGSGTISAFGTAKNAITVGAET----- 185
 473 AVDELTEKVGAVFVIAAGNPGVITVSGVATKXITVGAAPVPLVNGVYVSOALGYPD 532
 186 ----NVAESPFGSIADNNHIAQFSSRGATDGRKIKDVNTPGFTILASASSIAPDSSFWA 241
 533 YGFGYFPAITNV-----RIAFSSRGPRIDGELKRVVAPSGYIGTSSLPWIGADF-- 585

QY 242 NNSKAVNGCTSMATPIYAGNTAQLREHFIRKRGITPKPSLIKALLIGAT-----DV 295
 DB 586 -----MSGTMAPVHSGVALLISG-ARKEGIYVNDIIRKYLEGATWLEGDPYT 636
 QY 296 GLGYPGSDQGMGRVTLDKSLNVAVYNEATLTTGQATVSPFOQAKPKLISVWTDAG 355
 DB 637 GQKTELDQGHGLVNTVTKSWEI-----LKAINTTLPIDVHMADKSV 678
 QY 356 STTASVTLVNDLDTVTPAPG-----QKYGVN 392
 DB 679 SDFAEYLGV-DVIRGLYARNSIPDIYEMHIKYGD 712
 RESULT 5
 EXPR_XANCP STANDARD; PRT; 580 AA.
 ID AC P2314;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Extracellular protease precursor (EC 3.4.21.-).
 GN XCC0851.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_Taxid=340;
 RN [1]
 RP SEQUENCE FROM N. A.
 RX MEDLINE=90251253; PubMed=2187155;
 RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Daniels M.J.;
 RT "A multipurpose broad host range cloning vector and its use to
 characterise an extracellular protease gene of Xanthomonas campestris
 pathovar campestris";
 RL Mol. Gen. Genet. 220:433-440(1990).
 RP [2]
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Camaran F., Cardoso J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Fomiglieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
 RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities";
 RL Nature 417:459-463(2002).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to peptidase family 58.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
 CC or send an email to license@ebi.ac.uk).
 DR EMBL; X51635; CA35962.1; -
 DR EMBL; AE012184; AA040166.1; -
 DR PIR; S11890; S11890.
 DR HSSP; P00782; 28BT.
 DR MEROPS; S08.TPA; -.

```

DR InterPro: IPR000209; Peptidase_S8.
DR InterPro: IPR007280; PPC.
DR InterPro: IPR009020; Protease_inhib.
DR Pfam: PF00082; Peptidase_S8; 1.
DR Pfam: PF04151; PPC; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR HydroLase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 32
FT PROPEP 33 32
FT CHAIN 7137 580
FT ACT_SITE 177 177
FT ACT_SITE 237 237
FT ACT_SITE 409 409
FT DISULFID 225 273
FT DISULFID 315 352
FT DISULFID 450 454
SQ SEQUENCE 580 AA; 57228 MM; 8C9A2CBE4E7F47CB CRC64;

Query Match 13.1%; Score 296.5; DB 1; Length 580;
Best Local Similarity 29.0%; Pred. No. 4.1e-12;
Matches 137; Conservative 52; Mismatches 148; Indels 135; Gaps 26;

QY 21 GGGVVAVVADTGL---DGRNDSMHEAFRGKITLVLGRTNNAND----- 64
DB 168 GSGTVAVVADTGLTSTADLANAILAGYDFISDATTARDGNGSDSNADGDMVAANECA 227
QY 65 -----FNGGTHVAGS-----VLGNAIINKMAFOANL-----VFQSIMDS 99
DB 228 GIPAASSSMHGTHVAGVAAVNTNTTGAAGVAVRYVLGKCGSLSDIADAIWA 287
QY 100 SGG-IGGLPBNLN--TLFQAMNAGARHTNSGAPVNAAYANRQVDYRANNDVYL 156
DB 288 SGGTVSGIPNANPAVINNSLGGGSCSTWQNA-INAAVSRGT-----TVV 334
QY 157 FAAGNEGPN-SGTISAPGTAKNAITGATENYRPSFGSIADPNHIAQPSRGATRDGRI 215
DB 335 VAAGNDASNVSG--SLPANCAVIAVAAT-----TSAGA-----KASYSNIGT----- 375
QY 216 KPDVTAAGTITLARS--LAPDSFWANYSKYAMGTSMTATPVAAGNAQLREHIK 273
DB 376 GIDVSAFGSSILSTLNSGTTTSGSASVSYN-----GTSMASPHVAAGVALVOS--VA 426
QY 274 NRGITPK--PSLIK--AALIAGATDVGLG-----PSGDQMGRTVLDSK 314
DB 427 PTLATPAAVETTLKNTARALPGACSGGCGAGIVNAATAALINGSGGGGGGNTLNG 486
QY 315 LNVAYVNEATALTGQKATYSFQTAGKPKISLVTPDAPGSTTASVTLVNDLVI--- 371
DB 487 TPVTGGA---TGAEIYNTITVPAAGS---TLVTTTSGS-----GADIYVPA 531
QY 372 TAPNGKXYVNDPSYPDNMNGNNVENVFINAPSGTYTTEVOAYNVPSG 423
DB 532 SAPTDAAYT---CRPYRS-----GNAETCTIAP-SGTYVRLKXAYSTPSG 573

```

```

RT "Structure of a new alkaline serine protease (M-protease) from
RT Bacillus sp. KSM-X16."
RL Acta Crystallogr. D 51:199-206(1995).
RN [2]
RP SEQUENCE OF 1-23, AND CHARACTERIZATION.
RX MEDLINE=95358832; PubMed=763397;
RA Kobayashi T., Hakamada Y., Adachi S., Hitomi J., Yoshinatsu T.,
RA Koike K., Kawai S., Ito S.;
RT "Purification and properties of an alkaline protease from
RT alkalophilic Bacillus sp. KSM-X16."
RL Appl. Microbiol. Biotechnol. 43:473-481(1995).
CC -1- COFACTOR: Binds 2 calcium ions per subunit.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC PDB; 1MPT; 22-JUN-94.
DR MEROPS; S08.010; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DX HydroLase; Serine protease; Metal-binding; Calcium-binding;
DX 3D-structure.
FT ACT_SITE 32 32
FT ACT_SITE 62 62
FT ACT_SITE 215 215
FT METAL 2 2
FT METAL 40 40
FT METAL 73 73
FT METAL 75 75
FT METAL 77 77
FT METAL 79 79
FT METAL 163 163
FT METAL 165 165
FT METAL 168 168
FT HELIX 6 10
FT TURN 11 12
FT TURN 13 18
FT TURN 19 20
FT TURN 24 25
FT TURN 27 32
FT TURN 39 40
FT TURN 43 48
FT TURN 51 52
FT HELIX 62 71
FT TURN 84 85
FT STRAND 87 92
FT TURN 96 97
FT HELIX 102 114
FT TURN 115 116
FT STRAND 119 122
FT STRAND 126 126
FT HELIX 131 142
FT TURN 143 144
FT STRAND 146 150
FT STRAND 161 161
FT TURN 162 164
FT TURN 166 167
FT STRAND 176 177
FT STRAND 180 180
FT TURN 182 183
FT TURN 188 189
FT STRAND 192 195
FT STRAND 199 203
FT TURN 204 206
FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249

```



```

FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249
FT HELIX 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26698 MW; 4D89F8778999BE8D CRC64;

Query Match 11.8%; Score 269; DB 1; Length 269;
Best Local Similarity 31.8%; Pred. No. 9.3e-11;
Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;

CY 13 AONNYGLYGGGVAVAVADTGLDGRNDSMHEAFKRTALVALGRTNNANDPNHGTHV 72
DB 15 AANNRGITSGVAVAVADTGLDST-----HPDLNIRGASFPVPSGP-STOGNGHGTHV 66
CY 73 AGSV--LGNALN-KGNAPQANLVFOSIMDSGGIGGLPSNLTILFQAMNAGARIHTNSM 129
DB 67 AGTIALNNISIGVIGVAPSAELVAVKLGASG--SGSVASIAQGLEMAGNNGHVALNLSL 124
CY 130 GAVVNGAYTANSPQVEYVRNNMTVLPAAGNGSPNSGTTISAPGTAKNAITGATENTRP 189
DB 125 GSP--SPSATLTQAVNSATSGVLVVAASGNSG--AGSISYFARYANMAVAGAT----- 174
CY 190 SFGSIALDPNPHIAQFSSRGATRDGRIPDVTAPGTFTILARSLSLADSSFMVANSKYAY 249
DB 175 -----DONNRRASFSQYGAGL-----DIVAPGVAVQSTYP-----GSTYAS 210
CY 250 MCGTSAATPVAVGNA-----QIREFTNRRGTTPEKSLIKALAGATDVL 297
DB 211 LNSTSATPVAVAGAAALVKQKNSWNSVQIRNH-LKN-----TATSLGSTNL-- 256
CY 298 GYPSG 302
DB 257 -YGSQ 260

RESULT 8
ELYA_BACAO
ID ELYA_BACAO STANDARD; PRT; 380 AA.
AC P27693;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus alcalophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI TaxID=1445;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-PB92;
RC MEDLINE=91262463; PubMed=2059048;
RA van der Laan U.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
RA Quax W.J.;
RT "Cloning, characterization, and multiple chromosomal integration of a
RL Appl. Environ. Microbiol. 57:901-909(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RX STRAIN-PB92;
RC MEDLINE=92390330; PubMed=1518788;
RA van der Laan J.C., Tepiyakov A.V., Keiders H., Kalk K.H., Misset O.,
RA Mulleners L.J.M., Dijkstra B.W.;
RT "Crystal structure of the high-alkaline serine protease PB92 from
RL Bacillus alcalophilus.";
RN [3]
RP Protein Eng. 5:405-411(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).

```

```

EX MEDLINE=93078250; PubMed=1447775;
RA Sobek H., Hecht H.-U., Aehle W., Schomburg D.;
RT "X-ray structure determination and comparison of two crystal forms of
RT a variant (Asn154Y) of the alkaline protease from Bacillus
RT alcalophilus refined at 1.85-A resolution.";
RL J. Mol. Biol. 228:108-117(1992).
RN [4]
RP STRUCTURE BY NMR OF 112-380.
RX STRAIN-PB92;
RC MEDLINE=97277237; PubMed=9115441;
RA Martin U.R., Mulder P.A., Karim-Nejad Y., van der Zwan J.,
RA Martini M., Schipper D., Boelens R.;
RT "The solution structure of serine protease PB92 from Bacillus
RT alcalophilus presents a rigid fold with a flexible substrate-binding
RT site.";
RL Structure 5:521-532(1997).
CC -1- CORFACTOR: Binds 2 calcium ions per subunit.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65086; AAA2212.1; -;
DR EMBL; A13739; CA01126.1; -;
DR PIR; A49778; A49778.
DR PDB; 1AH2; 1S-APR-98.
DR MEROPS; S08.038; -;
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR000202; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00135; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR KW Hydrolyase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
KW Signal; 3D-structure.
FT SIGNAL 1 27
FT PROPEP 28 112
FT CHAIN 113 380
FT ACT_SITE 143 143
FT ACT_SITE 173 173
FT ACT_SITE 326 326
FT METAL 113 113
FT METAL 151 151
FT METAL 184 184
FT METAL 186 186
FT METAL 188 188
FT METAL 190 190
FT METAL 274 274
FT METAL 276 276
FT METAL 279 279
FT HELIX 117 120
FT TURN 121 123
FT HELIX 125 130
FT TURN 131 131
FT TURN 133 136
FT STRAND 138 142
FT TURN 150 151
FT STRAND 156 157
FT HELIX 173 182
FT STRAND 198 201
FT TURN 207 208
FT HELIX 213 226
FT STRAND 230 233
FT HELIX 242 252
FT TURN 253 255
FT STRAND 257 258

```

	FT	TUEN	267	268
FT	TUEN	267	268	
FT	TUEN	272	275	
FT	STRAND	279	279	
FT	STRAND	282	285	
FT	STRAND	291	291	
FT	TUEN	299	300	
FT	STRAND	303	306	
FT	STRAND	310	314	
FT	TUEN	315	317	
FT	STRAND	318	322	
FT	HELIX	325	342	
FT	TUEN	344	345	
FT	HELIX	348	358	
FT	HELIX	375	380	
SO	SEQUENCE	380 AA;	38853 NM;	539EA72771B6682C CRC64;
Query Match		11.9%;	Score 269;	DB 1; Length 380;
Best Local Similarity		31.8%;	Pred. No. 1.4e-10;	
Matches		97;	Conservative	35; Mismatches 99; Indels 74; Gaps 14
Db	13	ACNNYGLYGGCGVAVADTGLDGTGRNDSMEHAEFGKITALVALGRTNNANDENGHTYV	72	
Db	126	AANRGLTGSVGVAVALDTGTST-----HEDLINRGASFPVGEP-STQDNGHGHTY	177	
Qy	73	AGSV--LGNALN-KGVAPQANLVFQSIIMDSGCGGLGPEPSNNTLTFQAMNAGARITNSM	129	
Db	178	AGTIALNNISIGVLGAPPAELLYAKVLCASG--SGSVSSLAQLEHAKNGNGHVALNEL	235	
Qy	130	GAPNGAYTANSKQVDEYVRNNDMTVLPAGNEGNSCTIAPGTAKNALTVCATENYRP	189	
Db	236	GSP--SPSATLEQAVNSATSRGVLVVAASGNSG--AGSISYPRARYANMAVAT-----	285	
Qy	190	SFGSIADNPPIHIAQFSSRGATRGRLKPEVTPAFGFTLSASSLPDSSFMANYSKAY	249	
Db	286	-----DQNNRASFQOYAGL-----DIYAPGNNYSTP-----GSTYAS	321	
Qy	250	MGTSMATPIVAGNVA-----QREHFQKRGITPKPSLIKALLAGATDGL	297	
Db	322	INGSMATPEVAGAAALVQKQNPWSMNQIRNH-LKN-----TATSLGNTNL--	367	
Qy	298	GYPG 302		
Db	368	-YGS 371		
RESULT 9				
ELYA_BACCS		STANDARD;	PRT;	380 AA.
AC	P41362;			
DT	01-NOV-1995 (Rel. 32, Created)			
RC	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Alkaline protease precursor (EC 3.4.21.-).			
OS	Bacillus clausii.			
CC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=79880;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;			
RX	MEDLINE=93043753; PubMed=1368952;			
RA	Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,			
RA	Aono R., Horikoshi K.,			
RT	"Molecular cloning, nucleotide sequence, and expression of the			
RT	structural gene for alkaline serine protease from alkaliphilic			
RT	Bacillus sp. 221."			
RL	Biosci. Biotechnol. Biochem. 56:1455-1460(1992).			
RN	[2]			
RP	SEQUENCE OF 112-129.			
RC	STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;			
RA	Horikoshi K.;			
RA	(In) Horikoshi K. (eds.);			
RL	Microorganisms in alkaline environments, pp.187-194, VCH,			
RL	Weinheim (1991).			

```

CC -1- COFOR: Binds 2 calcium ions per subunit (by similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ch/announce/
CC or send an email to license@ebi.ebi.ch).
CC -----
DR EMBL, S48754; AAC60420.1; -
DR EMBL, D13157; BA02442.1; -
DR EMBL, A26817; CAA01836.1; -
DR EMBL, A22550; CAA01611.1; -
DR HSSP; P29600; IGTI.
DR MEROPS; S08.103; -
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR005020; Protease_fam1b.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KM Hydrolyse; Serine protease; Zymogen; Metal-binding; Calcium-binding;
KV Signal.
FT Signal. 1 27
FT PROPEP 28 111
FT CHAIN 112 380
FT ACT SITE 143 143
FT ACT SITE 173 173
FT ACT SITE 326 326
FT METAL 113 113
FT METAL 151 151
FT METAL 184 184
FT METAL 186 186
FT METAL 188 188
FT METAL 190 190
FT METAL 274 274
FT METAL 276 276
FT METAL 279 279
SQ SEQUENCE 380 AA; 38826 MW; 5F73ABCE8D5B6831 CRC64;
Query Match 11.9%; Score 269; DB 1; Length 380;
Best Local Similarity 31.88; Pred. No. 1.4e-10;
Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;
13 AQNNGYLYGQGVVAADTG;LDITGRNDSWMEAFRGKITALYALGRINNANDPNGHGTHV 72
126 AAHRRGITGSGVKAVALDTGIST-----HPDININGGASFFVGEPP-STQDGGHGTHV 177
73 AGSV--LGNALN-KGMAPOANLVQSIMDSGGGLPSNLNTLFSQAWNAGARIHTNSW 129
178 AGTIALNNTSGVUGVAPSAELVAVKVLGASG--SGSVSSTAGQLEWAGNNGHVAHLST 235
130 GAVPAGVATANSRQVDEVYRNNDMTVLFAANGEGPNSGTTISAPGTAKAALTVGATENYRP 189
236 GSP---SPSATLEQVANSATSRGLVVAASNGS--AGSISYPRAYNANMAVGAT----- 285
190 SFGSIADPNPIIAOFFSRGATRDGRIRKIPDVTAPEFTILASASSLAPDQSFWANTNSKYAY 249
286 -----DONNRBAFSQYAGL-----DIYAPGVNTQSTP-----GSTYAS 321
250 MGTSMATPIYAGNVA-----QLREHFKNRGITPKESLIKALLIAGATDVGL 297
322 LINGSMATPHVAGAAALVKQKPSMSVNVQIRNH-LKN-----TASTIGSTYL-- 367

```

OY 298 GYPSG 302
DB 368 -YSGG 371

RESULT 10

SUBT_BACS9
ID SUBT_BACS9 STANDARD; PRT; 420 AA.
AC P28842;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin precursor (EC 3.4.21.62).
GN SUBT.
OS Bacillus sp. (strain TA39).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29336;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9256481; PubMed=1581352;
RA Narinx E., Davail S., Feller G., Gerday C.;
RT "Nucleotide and derived amino acid sequence of the subtilisin from
the Antarctic psychrotroph Bacillus TA39.";
RL Biochim. Biophys. Acta 1131:111-113 (1992).
CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
it catalyzes the hydrolysis of proteins with broad specificity
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large unchanged residue
CC in P1. Hydrolyzes peptide amides.
CC -1- COFACTOR: Binds 1 calcium ion per subunit (potential).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: Still active at temperatures close to 0 degrees
CC Celsius; it has a marked heat lability.
CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
CC sporulation, and many mutations which block sporulation at early
CC stages affect expression levels of subtilisin. However, subtilisin
CC is not necessary for normal sporulation.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X62369; CAA44227.1; -
DR PIR: S23407; S23407.
DR HSSP: Q99405; 1MPT.
DR MEROPS: S08.004; 1MPT.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILISIN_ASP; 1.
DR PROSITE: PS00137; SUBTILISIN_HIS; 1.
DR PROSITE: PS00138; SUBTILISIN_SER; 1.
KW Hydrolyase; sporulation; Serine protease; Zymogen; Metal-binding;
KW Calcium-binding; Signal.
FT SIGNAL 1
FT PROPEP 27
FT CHAIN 112
FT ACT SITE 145
FT ACT SITE 182
FT ACT SITE 360
FT METAL 115
FT METAL 154
SQ SEQUENCE 420 AA; 44086 MW; AE4F12BD3B26EC CRC64;
Query Match 11.7%; Score 265.5; DB 1; Length 420;
Best Match Similarity 29.0%; Pred. No. 2, 8e-10;
Matches 92; Conservative 46; Mismatches 120; Indels 59; Gaps 15;

OY 21 GGGVAVADTGLDTGRNDSMEAFRGKITAL--YALGRT---NNANDPNHGTHVAGS 75
DB 136 GGGINIAVLDTGVTN-----HPDLRNNEVGCKDFVTGTTNNNCTBRQGHGTHVAGS 189
OY 76 VL-----GNALNKGMAQALVFOSTV--DSSGLGLGSPENLNTLFSQANNAARLHTN- 127
DB 190 ALADGTGNGV-YGVAADADLMAYKVLGDGSGYADDIAAIIHADQATALTQKVINM 248
OY 128 SWGAPVNGAYTANSRCQYDEVYRNNDMTVLPAAENEGNSGTTISAPGTAKNAITVGATENY 187
DB 249 SLSSGSSSLITNA--VNSYKGVLIILAAAGNSGPYQSGICYPGALVNAVAVALLEN- 304
OY 188 RPSFGSIADNPNNHIAQSSRFG-----ATRDGRKPDVTAPGTTILSARSLAPDSSF 239
DB 305 ----KYENGTYRVADFSRGSYMTGDVYIQKGDV--ELISAGAAIYST----- 347
OY 240 WANYNSKYAVMGSGMATP-VAGNVAQLREHFKNRGITPKPLIKALIIAGATDVLGY 299
DB 348 W--FDGGYATISGTSMAAPHAAGLAKIMAQYPSASNVDRSELQYRAY---ENDILSGY 402
OY 300 PSG-----DQWGRVTL 311
DB 403 YAGYGDDPFAAGFGFATV 419

RESULT 11

PROA_VIBAL
ID PROA_VIBAL STANDARD; PRT; 534 AA.
AC P16588;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alkaline serine exoprotease A precursor (EC 3.4.21.-).
GN PROA.
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=663;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83326126; PubMed=2546861;
RA Deane S.M., Robb F.T., Robb S.M., Woods D.R.;
RT "Nucleotide sequence of the Vibrio alginolyticus calcium-dependent,
RT detergent-resistant alkaline serine exoprotease A.";
RL Gene 76:281-288 (1989).
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M25499; AAA27550.1; -
DR PIR: JS0173; JS0173.
DR HSSP: Q99405; 1MPT.
DR MEROPS: S08.050; 1.
DR InterPro: IPR000209; Peptidase_S8.
DR InterPro: IPR007280; PPC.
DR InterPro: IPR009020; Protease_inhib.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILISIN_ASP; 1.
DR PROSITE: PS00137; SUBTILISIN_HIS; 1.
DR PROSITE: PS00138; SUBTILISIN_SER; 1.
KW Hydrolyase; Serine protease; Zymogen; Signal.
FT SIGNAL 1
FT PROPEP 22
FT CHAIN 142
SQ SEQUENCE 534 AA; 534 MW; ALKALINE SERINE EXOPROTEASE A.

FT ACT_SITE 180 180 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 363 363 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 534 AA; 55930 MW; 84E96DC649DA226 CRC64;

Query Match 11.7%; Score 265; DB 1; Length 534;
Best Local Similarity 26.3%; Pred. No. 4, 1e-10; Indels 128; Gaps 24;
Matches 123; Conservative 65; Mismatches 152;

QY 7 IVKADVAQ-----NNY--GLYGGQGVAVADTGLDTRGDSMHEAF 46
DB 137 IVSADANQNMALWGLDRIDQRLPLDNNNSANPDDGVTAVYIDGV-----NNAHAF 190
QY 47 RGTITLY-ALGRTNNANDPNNGHTVAGSVLGNALNKMMAQANLVQSITM--DSGGL 103
DB 191 GGRSVSGYDFVNDADASDCNGHGHVAGTIGSILY--GVAKNVNLVGVRLVSCSGSGST 248
QY 104 GGLPSNLNTLPSQANAGARHTNSWGAPEVNGAYTANSRQVDEYRNNDMTVLFAAGNEG 163
DB 249 SGVIAGDVVAANA--SGPSVANMSIG---GGQVALDASAVOSAVQSG-VSTMLAAGNSN 302
QY 164 PMSGTISAPGTAKNAITVGATENYRPSFGSIADNPNHIAQSSRGATRDGRIKPDVYAPG 223
DB 303 ADACNYS-PARVATGVTVGST-----TSTDARSSFSNMWQSC-----VDVFAPG 344
QY 224 TFLSARSSLAPDSSFMANNYSKYAMGGMATPIVAGNVQLREHFKNRGITPKPSL 283
DB 345 SQTKSA-----W--YDGGYKTSISGSMATPRVAG-VAAI--YLOENSVS--PSQ 387
QY 284 IKAALIAA-----TDVGLGYPSG-----DOGMGRVTLDKSLNAVY 319
DB 388 VEALIVSRASTGKVTDRGSVNKLYSLTDADCGQCPDPTPPECKLTSGVPSV--- 444
QY 320 VNEATLTLTGQKATYFQQAQKELKLSLWTDAPSTASTATVINDLIVITANQGX 379
DB 445 ---GLSSGSGQVAYVYDVEAGQRLTVQW-----YGGSGDADLYRFP--GAK- 486
QY 380 VGNDFSYPDNNWDR-----NNVENVFINAPOSGLTIEVQAANVPSG 423
DB 487 -----FTLNAMDGRPFYKGNNETCTVATOSGRHVMIQGYSNVSG 527

RESULT 12
SUBB_BACLE STANDARD; PRT; 269 AA.
ID SUBB_BACLE STANDARD; PRT; 269 AA.
AC P29539;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin B1 (EC 3.4.21.62) (Alkaline protease).
OS Bacillus lentus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1467;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
RA MEDLINE=93085738; PubMed=1453465;
RA Godecette D.W., Paech C., Yang S.S., Melenz J.R., Bystroff C.,
RA Wilke M.E., Fletterick R.J.;
RT "The crystal structure of the Bacillus lentus alkaline protease,
RT subtilisin B1, at 1.4-A resolution."
RT J. Mol. Biol. 228:580-595 (1992).
CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
CC it catalyzes the hydrolysis of proteins with broad specificity
CC -1- CATALYTIC ACTIVITY: Hydrolysis of peptide amides
CC in pI. Hydrolyzes peptide amides.
CC -1- COFACTOR: Binds 2 calcium ions per subunit.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
CC sporulation, and many mutations which block sporulation at early
CC stages affect expression levels of subtilisin. However, subtilisin
CC is not necessary for normal sporulation.
CC -1- SIMILARITY: Belongs to peptidase family 58.

DR PDB; 1ST3; 31-JAN-94.
DR MEROPS; S08.003; -.
DR InterPro; IPR00209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR Hydrolyase; Sporulation; Serine protease; Metal-binding;
KW Calcium-binding; 3D-structure.
FT ACT_SITE 32 32 CHARGE RELAY SYSTEM.
FT ACT_SITE 62 62 CHARGE RELAY SYSTEM.
FT ACT_SITE 215 215 CHARGE RELAY SYSTEM.
FT METAL 2 2 CALCIUM 1.
FT METAL 40 40 CALCIUM 1.
FT METAL 73 73 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 75 75 CALCIUM 1.
FT METAL 77 77 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 79 79 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 163 163 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT METAL 165 165 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT METAL 168 168 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT METAL 6 6 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT METAL 11 12 CALCIUM 1.
FT METAL 13 18 CALCIUM 1.
FT METAL 19 20 CALCIUM 1.
FT METAL 24 25 CALCIUM 1.
FT METAL 27 32 CALCIUM 1.
FT METAL 39 40 CALCIUM 1.
FT METAL 43 48 CALCIUM 1.
FT METAL 51 52 CALCIUM 1.
FT METAL 62 71 CALCIUM 1.
FT METAL 84 85 CALCIUM 1.
FT METAL 87 92 CALCIUM 1.
FT METAL 96 97 CALCIUM 1.
FT METAL 102 114 CALCIUM 1.
FT METAL 115 116 CALCIUM 1.
FT METAL 119 122 CALCIUM 1.
FT METAL 126 126 CALCIUM 1.
FT METAL 131 142 CALCIUM 1.
FT METAL 143 144 CALCIUM 1.
FT METAL 146 150 CALCIUM 1.
FT METAL 161 161 CALCIUM 1.
FT METAL 162 164 CALCIUM 1.
FT METAL 166 167 CALCIUM 1.
FT METAL 176 177 CALCIUM 1.
FT METAL 180 180 CALCIUM 1.
FT METAL 182 183 CALCIUM 1.
FT METAL 188 189 CALCIUM 1.
FT METAL 192 195 CALCIUM 1.
FT METAL 199 203 CALCIUM 1.
FT METAL 204 206 CALCIUM 1.
FT METAL 207 211 CALCIUM 1.
FT METAL 214 231 CALCIUM 1.
FT METAL 233 234 CALCIUM 1.
FT METAL 237 246 CALCIUM 1.
FT METAL 247 247 CALCIUM 1.
FT METAL 249 249 CALCIUM 1.
FT METAL 254 257 CALCIUM 1.
FT METAL 258 259 CALCIUM 1.
FT METAL 261 261 CALCIUM 1.
FT METAL 264 267 CALCIUM 1.
FT METAL 268 269 CALCIUM 1.
SQ SEQUENCE 269 AA; 26623 MW; E8AF1A6A9E2676B CRC64;

Query Match 11.4%; Score 259; DB 1; Length 269;
Best Local Similarity 31.5%; Pred. No. 4, 1e-10;
Matches 96; Conservative 34; Mismatches 101; Indels 74; Gaps 14;

QY 13 AQNNYGLYGGQGVAVADTGLDTRGDSMHEAFRGKITLYALGRTNNANDPNNGHTHY 72
DB 15 AAHNRGLTSGSVKAVADTGLIST-----HPLNTRGGASFPVGP-STDGNGHGHIV 66

```
QY 73 AGSV--LGNALN-KOMAPQANLVFQSIWDSGGLGGLPSNINTLPSQAWNAGARIHTNSW 129
DB 67 AGTIALNNSIGVLGAPSAELYAKVLGADG--RGAISIAQGLEWAGNNGHVAHLSL 124
QY 130 GAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPGSGTISAPGTAKNAITVGTATENYRP 189
DB 125 GSP---SPSATLEQAVNSATRGVLVVAASGNSGASS--ISYPARYANAVAGAT----- 174
QY 190 SFGSIADPNHIAFPSSRGATDRGRIKPDVTAPGTFTLSARSLAPDSSFWANYSKYAY 249
DB 175 -----DONNRRASFQYAGL-----DIVAPGVNVQSTYP-----GSTYAS 210
QY 250 MGGTSMATPIVAGNVA-----OLREHFTKRGITPKPSLIKAALIAGATDVL 297
DB 211 LNGTSMATPHVAGAAALVKQNPNSNVQIRNH-LKN-----TATSLGSTNL-- 256
QY 298 GYPSG 302
DB 257 -YGS 260

RESULT 13
THES_BACSP STANDARD; PRT; 401 AA.
AC Q45670;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thermophilic serine proteinase precursor (EC 3.4.21.-) (Ak.1
DE protease).
OS Bacillus sp. (strain AK1).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95085262; PubMed=7993087;
RA Maciver B., McHale R.H., Saul D.J., Bergquist P.L.;
RT "Cloning and sequencing of a serine proteinase gene from a
RT thermophilic Bacillus species and its expression in Escherichia
RT coli.";
RL Appl. Environ. Microbiol. 60:3981-3988(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
RX MEDLINE=20057863; PubMed=10589904;
RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
RT "Calcium-mediated thermostability in the subtilisin superfamily: the
RT crystal structure of Bacillus Ak.1 protease at 1.8-A resolution.";
RL J. Mol. Biol. 294:1027-1040(1999).
CC -! COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.
CC -! SUBCELLULAR LOCATION: Secreted.
CC -! MISCELLANEOUS: Has a pH optimum of 8.5, a temperature optimum of
CC 75 degrees Celsius.
CC -! SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L29506; AAA63688.1; -.
CC PIR; I39974; I39974.
CC PDB; 1DBI; 18-NOV-99.
CC MEROPS; S08.009; -.
CC InterPro; IPR000209; Peptidase_S8.
CC InterPro; IPR009020; Protease_inhib.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00136; SUBTILASE ASP; 1.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
FT SIGNAL; 3D-structure.
FT PROPEP 25 121
FT CHAIN 122 401
FT ACT_SITE 160 160
FT ACT_SITE 193 193
FT ACT_SITE 347 347
FT METAL 126 126
FT METAL 168 168
FT METAL 169 169
FT METAL 171 171
FT METAL 179 179
FT METAL 184 184
FT METAL 186 186
FT METAL 204 204
FT METAL 207 207
FT METAL 209 209
FT METAL 211 211
FT METAL 297 297
FT METAL 300 300
FT METAL 323 323
FT DISULFID 258 258
FT TURN 127 128
FT HELIX 129 132
FT TURN 135 136
FT HELIX 137 139
FT TURN 140 140
FT HELIX 141 144
FT TURN 145 147
FT TURN 152 153
FT TURN 155 160
FT TURN 165 166
FT TURN 168 173
FT STRAND 174 179
FT TURN 180 183
FT STRAND 184 184
FT HELIX 193 202
FT STRAND 219 224
FT TURN 228 229
FT HELIX 234 246
FT TURN 247 254
FT STRAND 251 254
FT HELIX 263 274
FT TURN 275 276
FT STRAND 278 282
FT STRAND 285 285
FT TURN 298 299
FT STRAND 301 306
FT TURN 308 309
FT STRAND 312 312
FT TURN 314 315
FT STRAND 316 316
FT TURN 320 321
FT STRAND 324 327
FT STRAND 331 335
FT TURN 336 338
FT STRAND 339 343
FT HELIX 346 362
FT HELIX 363 364
FT TURN 367 376
FT TURN 377 377
FT STRAND 379 379
FT TURN 382 383
FT STRAND 384 384
FT TURN 385 387
FT STRAND 388 388
FT STRAND 391 392
FT HELIX 395 399
FT TURN 400 401
SQ SEQUENCE 401 AA; 42835 MW; 1C736EFA4A89F256F CRC64;
```


Query Match 11.3%; Score 255; DB 1; Length 401;
Best Local Similarity 31.7%; Pred. No. 1.3e-09;
Matches 85; Conservative 26; Mismatches 99; Indels 59; Gaps 11;

QY 14 QNNYGLY-----GOGQVAVADTGLDGRNDSMHEAFRGKITALYALGRTNNANDP- 65
DB 137 QNTYTDYANDVTGSSGQEIADVDTGVD-----YHPDLGKVINGYDF--VDNDYDPM 188

QY 66 --NGHOTHVAGSVLGNALN-----KGMAPQANLVFQSIMDSSGGLGSLPNLNTLFSQAWN 119
DB 189 DLNHHGTHVAGIAAETNNATGAGMAPNTRILAVRALDRNG--SGTILSDIADAIYAAD 246

QY 120 AGARIHTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGKSGTISAPGTAKNAI 179
DB 247 SGAEVINLSGC---DCHTTTLENVNAVNWKGSVVVLAAGNNG--STTFEPASVENVI 301

QY 180 TVGATENYRPSFGSIADPNPHIAQFSRSGATRDGRKPDVTFAPGTFLSARSSLAPDSSF 239
DB 302 AVGAVDQY-----DRLASFNSYGTW-----VDVAPGVDIVSTIT----- 336

QY 240 WANYNSKYAMGTSTVATPIVAGVLAQL 267
DB 337 ----GNRYAYMSGTMSASPHVAGLAALL 360

RESULT 14
SCAL_STRPY STANDARD; PRT; 1167 AA.

AC P15926;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C5A peptidase precursor (EC 3.4.21.-) (SCP).
GN SCA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.; AND SEQUENCE OF 32-41.
RX MEDLINE=90153964; PubMed=2406246;
RA Chen C.C., Cleary P.P.;
RT "Complete nucleotide sequence of the streptococcal C5a peptidase gene of Streptococcus pyogenes.";
RL J. Biol. Chem. 265:3161-3167(1990).
CC C-!- FUNCTION: THIS VIRULENCE FACTOR OF S.PYOGENES SPECIFICALLY CLEAVES THE HUMAN SERUM CHEMOTAXIN CSA AT LYS(68)-ASP(69) BOND NEAR ITS C-TERMINUS, DESTROYING ITS ABILITY TO SERVE AS A CHEMOATTRACTANT.
CC C-!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).
CC C-!- SIMILARITY: Belongs to peptidase family S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; J05229; AAA26960.1; .
CC DR PIR; A35066; A35066.
CC DR HSP; Q45670; 1DBI.
CC DR MEROPS; S08.020;
CC DR InterPro; IPR001899; Gram_pos_anchor.
CC DR InterPro; IPR003137; PA.
CC DR InterPro; IPR00209; Peptidase_S8.
CC DR Pfam; PF00746; Gram_pos_anchor; 1.
CC DR Pfam; PF02225; PA; 1.
CC DR Pfam; PF00982; Peptidase_S8; 1.
CC DR PRINTS; PR00723; SUBTILISIN.
CC DR PROSITE; PSS0847; GRAM_POS_ANCHORING; 1.

DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Repeat; Signal.
FT SIGNAL 1 31
FT CHAIN 32 1130
FT PROPEP 1131 1167
FT ACT_SITE 130 130
FT ACT_SITE 193 193
FT ACT_SITE 512 512
FT DOMAIN 1029 1104
FT DOMAIN 1034 1101
FT REPEAT 1034 1050
FT REPEAT 1051 1067
FT REPEAT 1068 1084
FT REPEAT 1085 1101
FT SITE 1127 1131
FT SITE 1130 1130
FT MOD_RES 1130 1130
SQ SEQUENCE 1167 AA; 128263 MW; D2DDCS2E5752DA5D CRC64;

Query Match 11.0%; Score 248.5; DB 1; Length 1167;
Best Local Similarity 20.7%; Pred. No. 1.3e-08;
Matches 127; Conservative 67; Mismatches 152; Indels 269; Gaps 25;

QY 21 GOGQVAVADTGLDGRNDSMHEAFR-----GK 49
DB 121 GAGTVVAVIDAGFDKN-----HEAWRLTDKTKARYQSKEDLEKAKKEHGITYGEWVNDK 174

QY 50 ITALYALGRTNNANDPNHGTHVAGSVLGNALN-----KGMAPQANLVFQSIMDSSG 101
DB 175 VAYHDYSKDGKTAQVDEGHGTHVAGSVLGNALN-----KGMAPQANLVFQSIMDSSG 101

QY 102 GLGGLPSNLNTLFSQAWNAGARIHTNSGAPVNGAYT-AN-----SRQVDEYVRNNDMTVL 156
DB 234 GLADYARNYAQAIRDAVNLGAKVINMSFG-----NAALAYANLPDETKAFDYAKSGVSIV 290

QY 157 PAAGNEG-----PNSGTISAPGTAKNAITVG----- 182
DB 291 TSAGNDSFGGKTRPLADHPDYGVVGTTPAADSSTLTVASYSPPKQLTETAMVKTDDQD 350

QY 183 -----ATENYRPS-----FGSIA----- 195
DB 351 KEMPVLSTNRPEPKAYDYAVANRGMKEDDFKVKGKIALIERGIDIDFKQKVANAKAGA 410

QY 196 -----DNP----- 200
DB 411 VGVLIYDNQDKGFPIELPNVDQMPAAFTSRKDKGLLLKDNPKOKTITFNATPKVLTASGTK 470

QY 201 IAOFSRSGATRDGRKPDVTFAPGTFLSARSSLAPDSSFWANYNSKYAYMGGTSMATPIV 260
DB 471 LSRFSSWGLTADGNIKPDIAAPGQDIL---SSVA-----NNKYAKLSGTSMASAPLV 518

QY 261 AGNVAQLREHF-INKRGITP--KPSLIKAALIAGNT-----DVGLGYPSGQGRVTLTD 312
DB 519 AGIMGLLOKQYQYTPDMTPSERLDLAKVLMSSATALYDEDEKAYFSPRQOGGAVDAK 578

QY 313 K-SLVAVYVNEATALTTGQK-----ATYSFQTAGKPLKI---SLVWTDAPGST 357
DB 579 KASATVYVTDKD--NTSSKVLNANVSDKFEVTVVHNKSDKPKQLYVQATVQTD----- 631

QY 358 TASYTLVNDLDELVITAPNGQKVGVNDPSYP-----YDNWDSGNVNVENFINAPQSGVTI 413
DB 632 -----KVDGKHFAIPKVLVEASW-----QKITIPANSRQVTV 665

QY 414 EVOAYNVSPGPORFS 428
DB 666 PIDA-----SRFS 673

RESULT 15
WERA_BACSU
ID_WERA_BACSU STANDARD; PRT; 894 AA.

AC P54423; O06726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell
DE wall-associated polypeptides CWB23 and CWB25].
GN WPRA OR BSU10770.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
RC STRAIN=168;
RX MEDLINE=97158234; PubMed=9004506;
RA Margot P., Karamata D.;
RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
RT growth, encodes a cell-wall-associated protease.";
RL Microbiology 142:3437-3444(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98015415; PubMed=9353931;
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
RT "Sequencing of regions downstream of addA (98 degrees) and citG (289
RT degrees) in Bacillus subtilis.";
RL Microbiology 143:3305-3308(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo J., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.X., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinol S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat R., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: NOT YET KNOWN; COULD BE INVOLVED IN PROTEOGLYCAN
CC DEGRADATION, BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
CC -!- SUBCELLULAR LOCATION: Cell-wall bridges.
CC -!- PTM: PROCESSED INTO CWB23 AND CWB25.
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
CC
CC EMBL; U58981; AAC25926.1; -;
DR EMBL; Y09476; CAA70641.1; -;
DR EMBL; Z99109; CAB12917.1; -;
DR PIR; F69730; F69730.
DR HSR2; Q45670; IDBI.
DR MEROPS; S08.004; -;
DR Subtilisin; BG11846; wprA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PRO0723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Serine protease; Cell wall; Zymogen; Signal;
KW Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 894 CELL WALL-ASSOCIATED PROTEASE.
FT CWPB23.
FT CHAIN 32 ?
FT PROPEP 2 413 POTENTIAL.
FT CHAIN 414 894 CWPB23.
FT ACT_SITE 466 466 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 650 650 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 9 9 V -> A (IN REF. 1).
FT CONFLICT 14 14 L -> I (IN REF. 1).
SQ SEQUENCE 894 AA; 96487 MW; 0F67C353B55P8DBC CRC64;
Query Match 10.9%; Score 246; DB 1; Length 894;
Best Local Similarity 23.6%; Pred. No. 1.4e-08;
Matches 103; Conservative 65; Mismatches 142; Indels 126; Gaps 17;
QY 25 VVAVADTGLDITGRNDSMSHEAFRGKITALYA---LGRTNNDPNHGTHVAGSVLGNAL 81
DB 457 LIAVDTGVDSTLAD-----LKGVRTDLGHNFVRNNAMDDQGHGVHAGIAAQSD 510
QY 82 N---KGMAPQANLVFQSMDSGGGLGPNLNTLFSQANNAGARIHTNSGAPVNGAY 137
DB 511 NGYSMTGLNAKAIIPVKVLDAG--SGDTBOIALGIKIYAADKGAIVNLSLG----GGY 564
QY 138 TANSRQVDYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAITTVGATENYRPSFGSIADN 197
DB 565 SRVLEFALYAADKVNLIASAASNDGENA--LSYPASSKYTVSVGAT-----NR 611
QY 198 PNHIAQFSSRGATRGRIKIPDVTAPGTFTLSARSLAPDSSFWANNYSKIYMGTSNAT 257
DB 612 MDMTADFSNYGKGL-----DISAPGSDI-----PSLVP-----NGNVTYMSGTSNAT 653
QY 258 PIVAGNVAVOLREHFIKNRGITPKPSLIKAALTAGATDVLGYPGSDQ----- 304
DB 654 PYAAAAGLL---FAQNPKL--KREVEDMLKKTADDISFESVDGGESELYDDYDPIEI 708
QY 305 -----GMGRVTLDSLVNA---YVNSATALT-----GOKATY 334
DB 709 PKTFPGVDHESGYGRLLNVKAVSAADLQKVNKLESTQTAVRGSAGEKGTLLIEMVNGKKLIG 768
QY 335 SFQTOAGKPLKSLYWDTPAGSTTASYTLVNDLDDVITAPNGSKYVGNDFSPYDNNWDG 394
DB 769 SAAGKDNKAFKNVIA-----TOKQDQVLYLKATG-----DA 800
QY 395 RNNVNVFINAQSGT 410
DB 801 KTSYKVVVWVGKPSGT 816

Search completed: March 31, 2004, 16:05:28
Job time : 8.49423 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 32.475 Seconds
(without alignments)
4206.909 Million cell updates/sec

Title: US-09-985-689A-3

Perfect score: 2263

Sequence: 1 NDVARGIVKADVAQNNGLY.....EVOAYNVSPQRFSLAIYH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25; *
1: sp_archea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mbc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2263	100.0	433	2 Q9AQR4	Q9Aqr4 bacillus sp
2	2251	99.5	433	2 Q9AQR1	Q9Aqr1 bacillus sp
3	2234	98.7	433	2 Q9AQR2	Q9Aqr2 bacillus sp
4	2020.5	89.3	434	2 Q9AQR3	Q9Aqr3 bacillus sp
5	1998.5	88.3	639	2 Q9AQR3	Q9Aqr3 bacillus sp
6	1994.5	88.1	640	2 Q9AQR3	Q9Aqr3 bacillus sp
7	468	20.7	1825	5 Q9AQR1	Q9Aqr1 dictyosteli
8	442.5	19.6	1702	5 Q9AQR1	Q9Aqr1 dictyosteli
9	422.5	18.7	654	17 Q9AQR9	Q9Aqr9 pyrococcus
10	399.5	17.7	561	16 Q9AQR2	Q9Aqr2 thermoaer
11	396.5	17.5	1239	16 Q9AQR2	Q9Aqr2 streptomyce
12	377.5	16.7	430	16 Q9AQR1	Q9Aqr1 oceanobacil
13	353	15.6	1253	16 Q9AQR6	Q9Aqr6 streptomyce
14	347.5	15.4	412	2 Q9AQR6	Q9Aqr6 thermoaer
15	346.5	15.3	1237	2 Q9AQR4	Q9Aqr4 streptomyce
16	345.5	15.3	412	16 Q9AQR6	Q9Aqr6 thermoaer

17	345.5	15.3	1102	2 P95694	P95694 streptomyce
18	344	15.2	444	16 Q9KBJ7	Q9Kbj7 bacillus ha
19	343.5	15.2	1105	2 Q8KX46	Q8Kx46 streptomyce
20	336	14.8	1220	16 Q9L0A0	Q9L0a0 streptomyce
21	334	14.8	891	1 Q93635	Q93635 thermococcu
22	329	14.5	824	2 Q45464	Q45464 bacillus sp
23	329	14.5	1208	16 Q82B14	Q82b14 streptomyce
24	327	14.4	1139	16 Q82139	Q82139 streptomyce
25	327	14.4	1245	16 Q9RL54	Q9rl54 streptomyce
26	321	14.2	435	16 Q8EMJ3	Q8emj3 oceanobacil
27	315.5	13.9	442	16 Q31788	Q31788 bacillus su
28	307	13.6	1398	1 Q9P9L1	Q9p9l1 pyrococcus
29	304.5	13.5	799	16 Q9KEM1	Q9kem1 bacillus ha
30	290.5	12.8	1135	1 Q9P9D1	Q9p9d1 uncultured
31	289.5	12.8	1407	16 Q816G4	Q816g4 bacillus ce
32	288	12.7	1345	1 Q54437	Q54437 staphylothe
33	284.5	12.6	959	16 Q8PWS7	Q8pws7 xanthomonas
34	282.5	12.5	1101	16 Q82CF0	Q82cf0 streptomyce
35	279	12.3	431	2 Q9S3L6	Q9s3l6 bacillus sp
36	277.5	12.3	1098	16 Q9L1Z8	Q9l1z8 streptomyce
37	272.5	12.0	1571	2 Q8GCW3	Q8gcw3 streptococc
38	269.5	11.9	715	2 P70765	P70765 alteromonas
39	269	11.9	434	2 Q54327	Q54327 bacillus sp
40	268	11.8	575	16 Q8PNW1	Q8pnw1 xanthomonas
41	266.5	11.8	1570	16 Q8E2V6	Q8e2v6 streptococc
42	266.5	11.8	1570	16 Q8DX06	Q8dx06 streptococc
43	266	11.8	530	2 Q8GBS2	Q8gbs2 vibrio sp.
44	266	11.8	586	16 Q8PAL8	Q8pal8 xanthomonas
45	265	11.7	966	16 Q8PB28	Q8pb28 xanthomonas

ALIGNMENTS

RESULT 1

Q9AQR4 PRELIMINARY; PRT; 433 AA.

AC Q9AQR4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROA.
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=127889;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D6;
RX MEDLINE=20568675; PubMed=1118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.,
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046402; BAB21265.1; -.
DR HSSP; Q45670; 1DBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004283; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON TER 1
FT NON TER 433 433
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;

```
Query Match      100.0%; Score 2263; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 2.4e-124;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
DB 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHNTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
DB 121 GARIHNTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIPKDVTPAGTFFLSARSSLAPDSSFW 240
DB 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIPKDVTPAGTFFLSARSSLAPDSSFW 240
QY 241 ANVNSKYAVMGTSMTATPVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300
DB 241 ANVNSKYAVMGTSMTATPVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300
QY 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLWVTDAPGSTTAS 360
DB 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLWVTDAPGSTTAS 360
QY 361 YTLVNDLDLVIITAPNGQKTVGNDPYPYDNNWGNVENVFINAPQSGTYTIEVQAYNV 420
DB 361 YTLVNDLDLVIITAPNGQKTVGNDPYPYDNNWGNVENVFINAPQSGTYTIEVQAYNV 420
QY 421 PSGPQRFSLAIVH 433
DB 421 PSGPQRFSLAIVH 433

RESULT 2
Q9AQR1
ID Q9AQR1 PRELIMINARY; PRT; 433 AA.
AC Q9AQR1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. SD521.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133780;
RN [1]_TaxID=133780;
RP SEQUENCE FROM N.A.
RC STRAIN=SD521;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeiki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.; enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046405; BAB21268.1; -.
DR HSSP; Q45670; 1DBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18F660DDC CRC64;

Query Match      99.5%; Score 2251; DB 2; Length 433;
Best Local Similarity 99.3%; Pred. No. 1.2e-123;
Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
DB 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHNTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
DB 121 GARIHNTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIPKDVTPAGTFFLSARSSLAPDSSFW 240
DB 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIPKDVTPAGTFFLSARSSLAPDSSFW 240
QY 241 ANVNSKYAVMGTSMTATPVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300
DB 241 ANVNSKYAVMGTSMTATPVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300
QY 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLWVTDAPGSTTAS 360
DB 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLWVTDAPGSTTAS 360
QY 361 YTLVNDLDLVIITAPNGQKTVGNDPYPYDNNWGNVENVFINAPQSGTYTIEVQAYNV 420
DB 361 YTLVNDLDLVIITAPNGQKTVGNDPYPYDNNWGNVENVFINAPQSGTYTIEVQAYNV 420
QY 421 PSGPQRFSLAIVH 433
DB 421 PSGPQRFSLAIVH 433

RESULT 3
Q9AQR2
ID Q9AQR2 PRELIMINARY; PRT; 433 AA.
AC Q9AQR2
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133779;
RN [1]_TaxID=133779;
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeiki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.; enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046404; BAB21267.1; -.
DR HSSP; Q45670; 1DBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
```

DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 433
 SQ SEQUENCE 433 AA; 45587 MW; B81291A803C775AE CRC64;

Query Match 98.7%; Score 2234; DB 2; Length 433;
 Best Local Similarity 98.6%; Pred. No. 1.2e-122;
 Matches 427; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIKADVQAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
 DB 1 NDVARGIKADVQAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60

QY 61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGWAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 120

QY 121 GARIHTNSGAPVNGAYTANSRQVDVYVRNNDMTVLPFAAGNEGPNSTGTSAPGTAKNAIT 180
 DB 121 GARIHTNSGAPVNGAYTANSRQVDVYVRNNDMTVLPFAAGNEGPNSTGTSAPGTAKNAIT 180

QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSLAPDSF 240
 DB 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSLAPDSF 240

QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGY 300
 DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGY 300

QY 301 SGOQWGRVTLDKSLNVAIVNEATALTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 360
 DB 301 NGDOQWGRVTLDKSLNVAIVNEATALTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 360

QY 361 YTLVNDLDELVTAPNGQKYGNDPSYPYDNNWGRNNVNFVINAPOSQGTIVIEVQAYNV 420
 DB 361 YTLVNDLDELVTAPNGQKYGNDPSYPYDNNWGRNNVNFVINAPOSQGTIVIEVQAYNV 420

QY 421 PSGPQRFSLAIHV 433
 DB 421 PSGPQRFSLAIHV 433

RESULT 4
 Q9AQR0
 ID Q9AQR0 PRELIMINARY; PRT; 434 AA.
 AC Q9AQR0;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROE.
 OS Bacillus sp. NV1.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=133781;
 RN [1]_TaxID=133781;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NV1;
 RX MEDLINE=20568675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046406; BAB21269.1; -.
 DR HSSP; P00782; 1SUP.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.

GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 434
 SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;

Query Match 89.3%; Score 2020.5; DB 2; Length 434;
 Best Local Similarity 88.2%; Pred. No. 3.5e-110;
 Matches 393; Conservative 28; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIKADVQAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
 DB 1 NDVARGIKADVQAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60

QY 61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 119
 DB 61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSIMDSSGGLGGLPSNVSTLFSQAYS 120

QY 120 AGARIHTNSGAPVNGAYTANSRQVDVYVRNNDMTVLPFAAGNEGPNSTGTSAPGTAKNAI 179
 DB 121 AGARIHTNSGAPVNGAYTANSRQVDVYVRNNDMTVLPFAAGNEGPNSTGTSAPGTAKNAI 180

QY 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSLAPDSF 239
 DB 181 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSLAPDSF 240

QY 240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGY 299
 DB 241 WANDSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGY 300

QY 300 PSGQWGRVTLDKSLNVAIVNEATALTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 359
 DB 301 PSGQWGRVTLDKSLNVAIVNEATALTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 360

QY 360 SVTLVNDLDELVTAPNGQKYGNDPSYPYDNNWGRNNVNFVINAPOSQGTIVIEVQAYNV 419
 DB 361 SVTLVNDLDELVTAPNGQKYGNDPSYPYDNNWGRNNVNFVINAPOSQGTIVIEVQAYNV 420

QY 420 VPSGPQRFSLAIHV 433
 DB 421 VPSGPQRFSLAIHV 434

RESULT 5
 Q9AQR3
 ID Q9AQR3 PRELIMINARY; PRT; 639 AA.
 AC Q9AQR3;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Protease.
 GN PROA.
 OS Bacillus sp. 9860.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=133778;
 RN [1]_TaxID=133778;
 RP SEQUENCE FROM N.A.
 RC STRAIN=9860;
 RX MEDLINE=20568675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046403; BAB21266.2; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.


```
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00664; ABC membrane; 1.
DR Pfam; PF00005; ABC tran; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC TRANSPORTER 1; 1.
DR PROSITE; PS00893; ABC TRANSPORTER 2; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR KW ATP-binding; Protease; Transport.
SQ SEQUENCE 1825 AA; 202641 MW; E28160BC78613A3B CRC64;

Query Match      20.7%; Score 468; DB 5; Length 1825;
Best Local Similarity 27.1%; Pred. No. 1e-18;
Matches 157; Conservative 78; Mismatches 167; Indels 178; Gaps 23;

QY 19 LYGGQVAVADTGLDGR-----NDS-----SMHEAFRGKITALYALGRNNDPNGH 68
DB 327 LRGRQILSTADTGLDGHCHFFSDSNPIPNYNLNRKXVTVIGSL--HDNEDYVDGH 384
QY 69 GTHVAGSVLGN-----ALN--KGMAPQANLVFQSI-WDSSGGLGLPSNLTLFQAWN 119
DB 385 CTHVCGAAGAPEDSSSLAISFSGLATDAKIAFPDLASDFSNNEPVPEDYSQIQLFLN 444
QY 120 AGARIHTNSMGA-----PVNGAYTANSRQVDEYVRRN--DMTVLFAAGNEGNSGTISAPGT 174
DB 445 AGARVHGDSMGLSGIQQVLYGSDAGSIDDFLTHPFDLILRAAGNNEQYSSLLS-QAT 503
QY 175 AKNAITVGA----- 193
DB 504 AKNIVTGAETHTSYTTDALEYNPFETVAKSTLNSLCOSFDDDKYCTYTTACCTEYST 563
QY 184 -----TENRPSFGSIAD-----NPNHIAOFFSRGATRDGRIKPDVTAPGTITLSA 229
DB 564 VKGLSGCCTSVIKNSYASIPSSQPELYNENNICSFSKGPTHDGLKPDIVAFQGITSA 623
QY 230 RSSLA-----PDSFWANYSKYAYMGGTSMATPIVAGNVAQLREH-----FIK 273
DB 624 RSGANTTDQCGDGL-PNTNALLS-ESGTSMATPLATAATTILRQVLVDGYPTGSIVE 691
QY 274 NRGITPKPSLIIKAALIAAGATDVGLGYPSGD-----GQWGRVTLDKS 314
DB 682 SNKIQPTGSLKALMINNAQLNGTFFPLSSNTNPNFNAVPTFAGANFVGWGLMSLSEW 741
QY 315 LNVAIVNEA-----TALTGQKATYSF-----QTQA 340
DB 742 L---YVSSGVKPKPSRWVGIGELGDKKXASNWKKEYSLSTQNVSYCYFTYKPSSSGNSG 798
QY 341 GKP-LKISLVWTDAPGSTTASYTLVNDLILVITAPNGQ-KYVGNDFSYPYDNN----- 391
DB 799 GIPRIVATLVWTDPPSYSGAKNLVNNLDLMTNTESEFIYNSGSGSSNGTKGTTPL 858
QY 392 WDGRRNVNVEF---INAPQSGTYTIEVQANVPSPQRF 428
DB 859 QDSINNVEGIITYPTNKSEISPRFIAGTNIPGPQNF 898

RESULT 8
Q9GTN7
ID Q9GTN7 PRELIMINARY; PRT; 1702 AA.
AC Q9GTN7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TAGA.
GN TAGA.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RA Good J.R., Cabral M., Kuspa A.;
RT "taga", a putative serine protease/ABC transporter of Dictyostelium
RT that is expressed at the onset of development and is required for the
RT differentiation of a subpopulation of presore cells.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AF263455; AAG11416.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00664; ABC membrane; 1.
DR Pfam; PF00005; ABC tran; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER 1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER 2; 1.
DR KW ATP-binding; Transport.
SQ SEQUENCE 1702 AA; 187103 MW; 4A67716303CB7131 CRC64;

Query Match      19.6%; Score 442.5; DB 5; Length 1702;
Best Local Similarity 26.1%; Pred. No. 2.8e-17;
Matches 142; Conservative 91; Mismatches 168; Indels 143; Gaps 22;

QY 18 GLYGQGVAVADTGLDGR---NDSSMHEAFRGKITALYALGRNNDPNGHGHVAG 74
DB 300 GIKGDEIVGCADTGIDINHCFFYDTNFIGTHRKIIIS-YSSGNGDQIDEDGHGHV 358
QY 75 SVLGNAL-----NKGMAPQANLVFQSIWSSGGLGLPSNLTLFQANVAGARIHTN 127
DB 359 TIIGTTVDPPSVSEFSGGAPNSKVAFLVQLVQSGNGLSIQSNLTAIYQSTYDQNAK 418
QY 128 SMGAPVNGAYTANSRQVDEYVRRN--DMTVLFAAGNE---GPNS-GTISAPGTAKNAITVG 182
DB 419 ANNSNIGFFYTCVTEIDRFQWDHPDFLVRSAGNNVNFGENSYTILSQESTSKNSLV 478
QY 183 ATENTRPSFGSIAD----- 197
DB 479 SSNPSSSTYLSIDYWDWDFIYNSIRTSVCTQGQSIYGITCSDVPTQTTSYDIQTQCC 538
QY 198 P-----NHIAQ-----FSSRGATRDGRIKPDVTAPGTIFLSARSSL 233
DB 539 PILAKICSTRIQQOYQNTSVYSEFIPSLFSGVGPTSDGLKPDLLAPGSPFISSR-SL 597
QY 234 APDSSFANY-----NSKIAYMGGTSMATPIVAGNVAQLRE-----HFIKNR----- 275
DB 598 GPSSTI--NHCSPITSGIATSGALIAEGSSQAAAVATSAVLVRQYVRDGYFNGKVNSS 655
QY 276 -GITPKPSLIIKAALIAAGAT---DVLGYPSGQWGRVTLDK-----SLNV-AYV 320
DB 656 VGFQPSASLVKATLINTASINVDSTLEY---SQFGNQLSKLITTTNAQTSLDIPSSI 712
QY 321 NEA-TALTGQKATYSFQTAGKPLKISLVWTDAPGSTTASYTLVNDLILVITAPNGQKY 379
DB 713 EKADPIINTGETNSYCFSLDSKADIDITLVWTDAPGSPSLSTFTLVNLDLALLA-----F 767
QY 380 VGNDFSYPYDNN-----WDGRNVNVEFINAPQSGTYTIEVQANVPSPQRF 427
DB 768 VDGELSI-YSGNSETIFKNTSQVIFDQLNNVEIRIKDAP-IGSYDVKIFGTNIPNQSY 826
DB 428 SLAI 431
DB 827 SVVI 830
```

```

RESULT 9
Q8UOC9 PRELIMINARY; PRT; 654 AA.
AC Q8UOC9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 25, Last annotation update)
DE Alkaline serine protease.
GN PF1670.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010265; AAL61794.1; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Protease, Complete proteome.
SQ SEQUENCE 654 AA; 70230 MW; 1CB145A5F505DB34 CRC64;

Query Match 18.7%; Score 422.5; DB 17; Length 654;
Best Local Similarity 31.4%; Pred. No. 1.2e-16;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGQGVAVADTGLDGTGRNDSMHEAFRGKITALYALGRNNAN-----DPNGHG 69
DB 152 NLGYDGSGITIGITDGTID-----ASHPDLQGV-----IGWVDFVNGRSYPYDDRHG 200
QY 70 THVAGSVLG-----NALNKGMAPOANLVFQSIM--DSGGGLGGLPSNLTLFSQAWNAGA 122
DB 201 TVASIAAGTGAASCKYKGNAPGAKLAGIKVLGADSGSGISTIKGVAVAVNDKDKGI 260
QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPQAKNAITVG 182
DB 261 KVINLSLGSQSSDGTDLALSQAVNAADAGLVVVVAGNSGPNKYTTIGSPAAASKVITVG 320
QY 183 ATENYRPSFGSIADPNHIAQFSSRGATRDGRIPKDYTPAGTFTILSARSLAPDSFWAN 242
DB 321 AVDKY-----DVITFSRGPTADGRLPKPEVVAPGNWIIAARAS---GTSMGQP 366
QY 243 YNSKIAYNGGTSMATPIVAGNVAQLREHFIRNGITP---KPSLIKAALIAGATDVG-LG 298
DB 367 INDXYTAAPGTSMATPHVAGIAALLQ---AHPSTWPKVKATLIETADIVKPEIDIA 423
QY 299 YPSGQDQGRVTLKSLNVAVYNEATALTGQKA-----TYSFQTQAGKPLKISLWTD 353
DB 424 Y-----GAGRNVAYKAIN--VDNYAKLVFTGVANKGSQTHQFVISGASFYATLYWDNA 476
QY 354 PGSTASTYTLVNDLVLITAPNGQKYGVNDPFSYPYDNNWDRNNVNFVINAPOSQGYTI 413
DB 477 N-----SDLDLYLDPNQV---DYSY-----TAYGFEKVGYNPTDGTWTI 517
QY 414 EVQAVNVSPGQRFSLAIV 432
DB 518 KTVSY---SGSANYQDVV 533

```

RESULT 10
Q8RBJ2

```

Q8RBJ2 PRELIMINARY; PRT; 561 AA.
AC Q8RBJ2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Subtilisin-like serine proteases.
GN APR22 OR TTE0824.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
  Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
  Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
DR EMBL; AB0113049; AAM24081.1; -.
DR GO; GO:0005503; F:calcium ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002029; Peptidase S8.
DR InterPro; IPR007280; PPC.
DR InterPro; IPR009020; Protease inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Protease, Complete proteome.
SQ SEQUENCE 561 AA; 59968 MW; BA9C5C52F7083A18 CRC64;

Query Match 17.7%; Score 389.5; DB 16; Length 561;
Best Local Similarity 30.8%; Pred. No. 2.1e-15;
Matches 141; Conservative 64; Mismatches 160; Indels 93; Gaps 19;

QY 6 GIVKADVAQNNYGLYGQGVAVADTGLDGTGRNDSMHEAFRGKITALYALGRNNAN-- 63
DB 155 GITK---AKSDFGVTKMTITAITDGTIDGNHVDLS-----GGKI-----IGWKDFINNK 201
QY 64 ----DPNGHGTHV-----AGSVLGNALNKGMAPOANLVFQSIMDSG-----GLGGL 106
DB 202 TTPYDDNGHGTHVASIAAGTGAGNSFYKGVAPDALLVGIKVLGDANGSGSMSTVTAGIDWA 261
QY 107 PSNLTLFSQAWNAGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGNS 166
DB 262 VQNKDVGKIVNLISLGTSTSSDG--TDSTSLAVNRAVD-----SGIVVVVAAAGNSGPAK 314
QY 167 GTISAPGAKNAITVAGTENYRPSFGSIADPNHIAQFSSRGATRDGRIPKDYTPAGTFTI 226
DB 315 YTIISGPAEAKAITVAAD---VGLGFN---LASFSSRGPTADGRIPKDIAPGNYI 367
QY 227 LSARSLAPDSFWANYSKYAYMGGTSMATPIVAGNVAQLREHFIRNGITPPELSLIKA 286
DB 368 TAAK-----ANSVNGYVYTSGTSMATPFVAGTVAGTVALMLN---ANPNLTNDNA--KN 412
QY 287 ALIAGATDVLGYPGSDQGWRTLDKSLNVA-----YNEATALTGQKAT 333
DB 413 IYNSAKSGPPSKVDYAGRLDGYEAIRVAGNFRGNNDVPNHYIISGLVPSRYSDT 472
QY 334 YSFQ-QQAGKPLKISLWTDAPGSTASTYTLVNDLVLITAPNGQKYGVNDPFSYPYDNNW 392
DB 473 WTFNATNYSYPIAITLIIPD-----WANYN--PDPDIYLDPSGTLIKSS----- 515
QY 393 DGRNNVNFVINAPOSQGYTIEVQAVNVSPGQRFSLA 430

```


219	DB	VEADLADSTAQIGAPRAVAGNGTGGVEVAVLDTGVDAG-----HPDLADRIAARQSFV	272
58	QY	RTNNANDPNGHGTHVAGSYLGNAL-----NKGMAPOANLVFQSIIMDSSGGLGGLPSNLNT	112
273	DB	PBENTDDRDGCHTHVASTIAGTGAASAGKEKGVAPCARLSIGKVLDNS-GRGQISWTLAA	331
113	QY	LFSQAWNAGARHTHSWGA-PVNGAYTANSROVDEYVRNNDMTVLPAAGNEGPNSGTISA	171
332	DB	MEWAVERHAKLVNNSLGSGEQSDGSDPMSRAVDRLSAQTALFVVAAGN-GGEAGSIGA	390
172	QY	PETAKNAITVGTATENYRPSFGSIANDPNHIAQFSSRGATRDGRIKPDTVAPGTFFILSARS	231
391	DB	PGVAISALTVGA-----VDATDTLAPSSQGPVDCALKEPITAPGVGILAA--	437
232	QY	SLAPDSGFWANTNSKYAVMGTSMTAPVIVAGNVAQLREHFKNRGITPKPSLIKAL---	288
438	DB	-----NSSFAAGNGAYQSLSGTSMATPHVAGAALL-----AAARPDLSGSALKDV	484
289	QY	IAGATDVLGYPGSGQCGWRVTLDKSLNVAVYNEATALTTGQK-----ATYSFOTQAG	341
485	DB	LASSHRTPRYDAFQAGSGRVDVDAVRACGVASATAYAPGSSPGVRRLVITYTNTGAA	544
342	QY	KPLKSLVWTDAPGST---TASYTLV---NDLDLVIT-----APNQKQYVGNDFSYPDN	390
545	DB	VTLELSVAATHAPEGVFRLSASRVTPAHGTADVTILTDGSGAGGRAYSQGILAA---T	600
391	QY	NWDGRNVENVINAPQSGTYTIEVQAYNVPSGPQRFSLAI	431
601	DB	DADARN-----VAHTAVSAGPVRHKUTV	623
RESULT 12			
Q8ENV1		PRELIMINARY; PRT; 430 AA.	
AC	Q8ENV1		
DC	Q8ENV1		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Intracellular alkaline serine proteinase.		
GN	OB2375.		
OC	Oceanobacillus ihovensis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.		
OX	NCBI_Taxid=182710;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=HTE831 / DSM 14371 / JCM 11309;		
RC	MEDLINE=2220767; PubMed=12235376;		
RA	Takami H., Takaki Y., Uchiyama I.;		
RT	Genome sequence of Oceanobacillus ihovensis isolated from the Iheya		
RT	Ridge and its unexpected adaptive capabilities to extreme		
RT	environments."		
RL	Nucleic Acids Res. 30:3927-3935 (2002).		
DR	EMBL; AP004601; BAC14331.1; --		
DR	GO; GO:0004289; F:subtilase activity; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR000209; Peptidase_S8.		
DR	Pfam; PF00082; Peptidase_S8; 1.		
DR	PRINTS; PR00723; SUBTILISIN.		
DR	PROSITE; PS00136; SUBTILASE ASP; 1.		
DR	PROSITE; PS00137; SUBTILASE_HIS; 1.		
DR	PROSITE; PS00138; SUBTILASE_SER; 1.		
KW	complete proteome.		
SQ	SEQUENCE 430 AA; 45938 MW; 6D09A999B8C1E310F CRC64;		
Query Match 16.7%; Score 377.5; DB 16; Length 430;			
Best Local Similarity 34.5%; Pred. No. 2.8e-14;			
Matches 119; Conservative 48; Mismatches 119; Indels 59; Gaps 16;			
QY	2	DVARGIVKADVQNNVGLYCGQGVAVADTGLDTRNDSMHAEFRKII--ALYALGRT	59
DB	121	DTASSINADVLES-GLTGGQSTIAVDITGHP-----HEDLEGRIGFADFVKGT	172
QY	60	NNANDPNGHGTHVAGSYLGN-ALN-----KGMAPQANLVFQSIIMDSSGGLGGLPS-----	108

Db 173 EYDD-NHGTTCAGDAAGSAGQYQGPAPDANLVGVYLNKGTG-GSLSTVIEGID 230
 QY 109 -----NLNLTFSQAWNAGAR-IHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLFAA 159
 Db 231 WCIONQSKYNIL--SLSGSDATEPAEGDPVNVAV-----ETAWNGMVVCVAA 279
 QY 160 NNEGNSGTISAPGTAKNAITVGATENVRPFGSTADPNHIAQFSSRGATRGRIKPDV 219
 Db 280 GNSGPGDKTVGSPGSPKVTITGAADDNNTAERS-----DDSVAFSSRGPTIDGLTKPIL 335
 QY 220 TAPGFFILSARS--SLAPDSSFWANYSKYAMVGTSMTATPVGAVNAQLREHFIKNGI 277
 Db 336 LPFGVDIVSLRAPGSFIDKTNKSARVGSYISLSTGTSNATPICAGIVAQLLQ---SDSSL 392
 QY 278 TPKPSLIIKAALAGATDVLGVPDQGWGRVTLDKSLNVAVNE 322
 Db 393 T--PNQVKELMEACQD--LQSPNVVAGYL---NAANLINE 430

RESULT 13
 Q9FC06 PRELIMINARY; PRT; 1253 AA.
 AC Q9FC06;
 DT 01-MAR-2001 (TremBLrel. 16, Created)
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE Putative secreted peptidase.
 GN SCO7176 OR SC8A11.04C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8943436;
 RA Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96 (1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Narraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147 (2002).
 DR EMBL; AL939130; CAC01576.1; --
 DR HSSP; Q99405; IMPT.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001317; PA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Complete proteome.
 SQ SEQUENCE 1253 AA; 130971 MW; AA69B417EFEDB89 CRC64;
 Query Match 15.6%; Score 353; DB 16; Length 1253;
 Best Local Similarity 30.1%; Pred. No. 3, 1e-12;
 Matches 143; Conservative 56; Mismatches 164; Indels 112; Gaps 25;
 QY 18 GYGQGVVAVADTGLDGTGRNDSMHEAFRGKITAYALGRNTNNANDPNHGTHVAGSVL 77
 Db 235 GNTGEGVAVLDTGVDA-----HPDPAGRIATAAGFVDPQDVTDRNGHGHVASTVA 288
 QY 78 G-----NALNKGMAPQANLVFQSIMDSSGGLGLFSLNLTLSQAW---NAGARHTNSW 129
 Db 289 GTGAASGGVKGVPAGASLHIGKVLDSG--SGQSWV--LAGMEWAVRDQHAKIVMSL 344
 QY 130 G-APVNGAYTANSROVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAITVGATENYR 188
 Db 345 GDSPTDGT-DPLSEAVNMLSAETGALFVVAAGNSGPEAYTGTTPAADAALTVGAVNG-- 401
 QY 189 PSFGSIADPNHIAQFSSRG-ATRDGRIPKPDVATGRTILSNRSSLPDSSFWANYSKY 247
 Db 402 PKG-----VDQLADFSRGRVGNVAKPDLTAPGVGLAARSYPAGE-----EGAY 450
 QY 248 AYMGGTSMATPIVAGNVAOL-REHFIKNGITPKP-----SLIKAALIAGATDVLGYPG 302
 Db 451 QSLSGTSMATPHVAGAAALLAAEH-----PDWTGQRLKEALV-GTTAGTORFSPF 499
 QY 303 DOGWGRVTLKSLNVAVNEATALTQO---KATYSF-----QTQAG-KPLKIS 347
 Db 500 DAGSGRV-----DVAAAVRSTLLASGDAFAQAHPYPTFGQTRDRDVTYTNSSGPAFALD 553
 QY 348 LVMTDA-----PGSTTASYTLVNDLD-----LVITAPNGQKYV 380
 Db 554 LALSAPLPEGLFTLSEQAQVTPAHGTASVGVITHLDAEDNGAVATLVSAGDAGVLA 613
 QY 381 GNDFSYPYDNNWDGRRNNVENVFINA-----PQSGTYTIEVQAVNVPSPGFSL 429
 Db 614 RT-----PVGKNKEGRR--ATLALTAKDHDKPFLSGTVILKDVERN--TAPKYYSV 660
 RESULT 14
 Q9AER6 PRELIMINARY; PRT; 412 AA.
 AC Q9AER6;
 DT 01-JUN-2001 (TremBLrel. 17, Created)
 DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE Thermicin (Subtilisin-like protease thermicin).
 OS Thermoanaerobacter yonsei.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.
 OX NCBI_TaxID=111519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jang H.J., Kim B.C., Pyun Y.R., Kim Y.S.;
 RT "A novel subtilisin-like serine protease from Thermoanaerobacter
 yonseiensis K8-1: cloning, expression and biochemical properties."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Jang H.J., Kim B.C., Pyun Y.R., Kim Y.S.;
 RT "Subtilisin-like protease, thermicin, from Thermoanaerobacter
 yonsei." (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

```

DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease.
SQ SEQUENCE      1237 AA;  129488 MW;  266C2B2760223D40 CRC64;

    Query Match               15.3%; Score 346.5; DB 2; Length 1237;
    Best Local Similarity   30.4%; Pred.No. 7.4e-12;
    Matches 140; Conservative 55; Mismatches 154; Indels 111; Gaps 23;

QY      6  GLVKADVAQN-----YELYGQGVVAVADTLDTGRNDSMHAFRGI--TAL 53
Db      213 GRVKADLADSTAQIGAKVWASGDTGDQVKVAMLDSGADT-----EHPDLVGQVSDFSAS 266

QY      54  YALGRTNANDNGEGTHVAGSVLG-----NALNKMAPQANLVFQSINDSSGLGLPS 108
Db      267 FVPGEDDIA-DYNGEGHVVASTIVGTGSADGKERGVASGARLSVGKVLNSEG--SGQBS 323

QY     109 NLNTLFQSAWNA--CARIHNSWGAPVNGAYTAN---SRQDEVYVRNNDMTVLFAAGNE 162
Db     324 WI--TAGMEWAARDQKARIISNLG----GGGKDNDPMSCAVDELSDHTGALPVIAANG 377

QY     163 GNSGTISIAPGAKNAITVGATENRPFSGSIADPNHIAQSPSRRG-AFRDGRHKPDVTA 231
Db     378 GPHS---ISSPGAADSALTVGA-----VDSTDTLADFSSQGPRDGDKGLKEITA 424

QY     222 PGTFTLSARSLAPDSSFWANYNSKYVMGTSMATPIVAGNVQAOLREHFINKRGITPKP 281
Db     425 PGVDIVAASHYKRSGSY-----YTMSGTSMATPHVAGVAALLAASHPDWTQTQKE 477

QY     282 SLIKAAIAGATDVGLGYPSGQGWRTVLDKSLNVAYVNEATALTGOKATYSFOFAQG 341
Db     478 ALVSSAKATPA-----YTPYQAGAGRLDAPAAVHTTVPATTAYS-----GFHT--- 521

QY     342 KPLKISLVMTDAPGST---TASYTLVND---LDLVI---TAPNGQKVGNDFESPYPDNKW 392
Db     522 -----WPPFGEDTRVTYTVNGDAPSVLNLAINGTVPAGL-----FSLSEDH-- 565

QY     393 DGRNNVENVFINAPOSGTYTIEV-QAYNVPSGPQRFLAI 431
Db     566 -----VTVPAHGATVTLTIALDKLAGDOSVASVI 595

```

Search completed: March 31, 2004, 16:08:57
Job time : 33.475 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.1304 Seconds
(without alignments)
2595.843 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQAYNVPSQGFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04: *
1: Geneseqp1980as: *
2: Geneseqp1990as: *
3: Geneseqp2000as: *
4: Geneseqp2001as: *
5: Geneseqp2002as: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004as: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2261	100.0	433	5	AAM50084 Bacillus
2	2251	99.6	433	5	AAM50082 Bacillus
3	2245	99.3	433	2	AAR26274 Alkali-pr
4	2245	99.3	433	2	AAR61495 Modified
5	2245	99.3	433	2	AAR95698 Bacillus
6	2245	99.3	433	3	AAY69207 Amino aci
7	2245	99.3	433	3	AAY44619 Bacillus
8	2240	99.1	433	5	AAM50083 Bacillus
9	2236	98.9	433	5	AAR89548 Bacillus
10	2044	90.4	433	5	AAM50086 Bacillus
11	2044	90.4	641	2	AAR89547 Bacillus
12	2024	89.5	434	5	AAM50085 Bacillus
13	2010	88.9	434	5	AAM50081 Bacillus
14	2002	88.6	639	2	AAY17089 Bacillus
15	1999	88.4	640	2	AAY17091 Bacillus
16	1998	88.4	434	5	AAM50080 Bacillus
17	1998	88.4	640	2	AAY17090 Bacillus
18	1952	86.4	639	2	AAY17087 An alkali
19	1952	86.4	640	2	AAY17088 An alkali
20	1841	81.4	434	5	AAM50090 Bacillus
21	1514	67.0	345	2	AAM62230 Subtilase
22	1514	67.0	345	2	AAY21654 Subtilase
23	443	19.6	659	2	Aaw24121 Thermococ
24	443	19.6	659	2	Aaw94840 WO9856926
25	422.5	18.7	412	2	Aaw94836 Hyperther

26	422.5	18.7	522	2	AAW24122	Aaw24122 Pyrococcus
27	422.5	18.7	522	2	AAW94838	Aaw94838 Hyperther
28	422.5	18.7	654	2	AAW24129	Aaw24129 Pyrococcus
29	422.5	18.7	654	2	AAW94841	Aaw94841 Hyperther
30	401.5	17.8	659	2	AAW24123	Aaw24123 Protease
31	373.5	16.5	545	4	ABSO9483	ABSO9483 T. yonsei
32	347.5	15.4	1079	6	AAW81180	AAW81180 Transglut
33	347.5	15.4	1079	6	ABU07391	ABU07391 Foreign p
34	343	15.2	520	2	AAW13666	AAW13666 Fragment
35	343	15.2	734	2	AAW13667	AAW13667 Streptomy
36	343	15.2	823	2	AAW13668	AAW13668 DhpA-mel
37	341.5	15.1	1237	6	ABU11343	ABU11343 Protein e
38	342	13.8	806	2	AAR27481	AAR27481 RP-III re
39	307	13.6	903	2	AAR87007	AAR87007 Hyperther
40	307	13.6	1398	2	AAR87008	AAR87008 Protease.
41	307	13.6	1398	2	AAW24124	AAW24124 Pyrococcus
42	307	13.6	1398	2	AAW94839	AAW94839 WO9856926
43	298.5	13.2	580	7	ADD24927	ADD24927 Xanthomon
44	295	13.0	519	6	ABP76735	ABP76735 Streptomy
45	295	13.0	19938	6	ABP76678	ABP76678 Streptomy

ALIGNMENTS

RESULT 1

AAM50084
ID AAM50084 standard; protein; 433 AA.

XX AAM50084;

AC AC (first entry)

DT 12-AUG-2002

DE Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.

KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

OS Bacillus sp.

PN EPI2092333-A2.

XX 29-MAY-2002.

PD 22-NOV-2001; 2001EP-00127851.

XX 22-NOV-2001; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.

XX (KAOS) KAO CORP.

FA Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

FI Okuda M, Saeki K;

XX WPI; 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions.

PT Claim 5; Page 16-18; 25pp; English.

FS This invention describes novel Bacillus sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased detergency (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This sequence represents a fragment of the alkaline protease SD-521 from Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the invention

XX Sequence 433 AA;

SQ Query Match 100.0%; Score 2261; DB 5; Length 433;

Best Local Similarity 100.0%; Pred. No. 3.8e-168;

Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 DB 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTGISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTGISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSLAPDSSEW 240
 DB 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSLAPDSSEW 240
 QY 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSLAPDSSEW 240
 DB 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSLAPDSSEW 240
 QY 241 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 SGDOGWGRVTLDKSLNVAVYNEATATGQKATYSFOAQAKPLKISLVWTDAPGSTTAS 360
 DB 301 SGDOGWGRVTLDKSLNVAVYNEATATGQKATYSFOAQAKPLKISLVWTDAPGSTTAS 360
 QY 361 YTLVNDLDELIVITAPNGQKVGNDFSYPYDNNWGRNNDVFNAPQSGTYYTIEVQAYNV 420
 DB 361 YTLVNDLDELIVITAPNGQKVGNDFSYPYDNNWGRNNDVFNAPQSGTYYTIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 2

AAW50082 standard; protein; 433 AA.

AAW50082;

12-AUG-2002 (first entry)

Bacillus sp D6-(FERM P1592) alkaline protease protein fragment.

Alkaline protease; detergent; laundry; bleaching; dishwasher.

Bacillus sp.

EP1209233-A2.

29-MAY-2002.

22-NOV-2001; 2001EP-00127851.

22-NOV-2000; 2000JP-00355166.

12-APR-2001; 2001JP-00114048.

(XAOS) KAO CORP.

Harada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

Okuda M, Saeki K;

WPI; 2002-437518/47.

New modified alkaline proteases useful in detergent compositions.

Claim 5; Page 13-15; 25pp; English.

This invention describes novel Bacillus sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased detergency (34 - 38) compared to prior art alkaline proteases (31 and 23). This sequence represents a fragment of the alkaline protease E-1 from Bacillus sp strain D6-(FERM-P1592) described in the method of the invention

XX Sequence 433 AA;
 SQ Query Match 99.6%; Score 2251; DB 5; Length 433;
 Best Local Similarity 99.3%; Pred. No. 2.3e-167;
 Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 DB 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTGISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTGISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSLAPDSSEW 240
 DB 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSLAPDSSEW 240
 QY 241 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 SGDOGWGRVTLDKSLNVAVYNEATATGQKATYSFOAQAKPLKISLVWTDAPGSTTAS 360
 DB 301 SGDOGWGRVTLDKSLNVAVYNEATATGQKATYSFOAQAKPLKISLVWTDAPGSTTAS 360
 QY 361 YTLVNDLDELIVITAPNGQKVGNDFSYPYDNNWGRNNDVFNAPQSGTYYTIEVQAYNV 420
 DB 361 YTLVNDLDELIVITAPNGQKVGNDFSYPYDNNWGRNNDVFNAPQSGTYYTIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 3

AAW50082 standard; protein; 433 AA.

AAW50082;

05-FEB-1993 (first entry)

Alkali-protease Ya enzyme.

Alkali resistance; surface active agent resistance; detergent improver.

Bacillus sp. Y.

JP04197182-A.

16-JUL-1992.

28-NOV-1990; 90JP-00327110.

28-NOV-1990; 90JP-00327110.

(LIOY) LION CORP.

WPI; 1992-288440/35.

N-FSDB; AAQ27516.

DNA coding alkali-protease Ya enzyme - has good alkali and surfactant resistance and improves detergency.

Claim 2; Page 1; 17pp; Japanese.

The sequence is that of alkali-protease Ya enzyme which can be used in

CC the recombinant production of Ya enzyme. Ya enzyme is excellent in alkali
 CC resistance and surface active agent resistance and improves detergent
 XX
 SQ Sequence 433 AA;

Query Match 99.3%; Score 2245; DB 2; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIPKPDVTAPGTFTLSARSSLAPDSFW 240
 DB 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIPKPDVTAPGTFTLSARSSLAPDSFW 240
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 SGQDQWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 DB 301 SGQDQWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 QY 361 YTLVNDLDLVTAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 DB 361 YTLVNDLDLVTAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 4
 AA61495
 ID AAW61495 standard; protein; 433 AA.
 AC AAW61495;
 XX
 DT 06-NOV-1998 (first entry)
 DE Modified Bacillus lion Y protease.
 KW Bacillus lion Y protease; polyethylene glycol; PEG; soap;
 KW methoxypolyethyleneglycol; mPEG; skin; hair care product; cosmetic;
 KW lipstick; hair gel; sun oil; shampoo; hair dye; insect repellent.
 OS Bacillus sp.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1..433
 FT /note= "The enzyme is modified by methoxypolyethyleneglycol
 FT molecules covalently attached to the N-terminal amino
 FT group and to fourteen unspecified amino groups of lysine
 FT residues present on the surface of the enzyme"
 XX
 FN WO9803682-A1.
 XX
 XX
 PD 16-JUL-1998.
 XX
 XX
 PF 12-JAN-1998; 98WO-DK000015.
 XX
 XX
 PR 10-JAN-1997; 97DK-00000038.
 PR 25-JUN-1997; 97DK-00000754.

XX
 PA (NOVO) NOVO-NORDISK AS.
 XX

PI Olsen AA, Pronto A;

XX WPI; 1998-399132/34.

XX
 DR
 XX

PT New enzyme modified by attachment of many polymeric molecules - useful in
 PT skin and hair care products, has reduced tendency to cause sensitisation
 PT and increased stability.

PS Claim 16; Page 44-45; 56pp; English.

XX The present sequence represents the Bacillus lion Y protease. The
 CC invention claims for enzymes covalently modified on their surface by the
 CC attachment of a large number of small polymeric molecules, e.g.
 CC polyethylene glycols (PEG). The polymeric molecules are coupled to the N-
 CC terminal amino group and the amino groups of lysine residues found on the
 CC surface of the enzyme. In the example given, the lion Y protease was
 CC modified using methoxypolyethyleneglycol (mPEG) as the polymeric molecule.
 CC The N-terminal amino group and the amino groups of the fourteen lysine
 CC residues present on the surface of the lion Y protease were modified.
 CC Modification of the enzymes increases the stability and/or reduces the
 CC sensitising potential (allergenicity) of the enzyme, without
 CC significantly reducing enzymatic activity. Also, using a large number of
 CC relatively small polymeric molecules, rather than a few very large ones,
 CC provides a more even effect with reduced activity loss. The modified
 CC enzymes are claimed to be useful as components of a wide range of skin
 CC and hair care products, e.g. soaps, cosmetics, creams, gels, lipsticks,
 CC hair gels, sun oils, shampoos, hair dyes, insect repellants, etc

SQ Sequence 433 AA;

Query Match 99.3%; Score 2245; DB 2; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIPKPDVTAPGTFTLSARSSLAPDSFW 240
 DB 181 VGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIPKPDVTAPGTFTLSARSSLAPDSFW 240
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 SGQDQWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 DB 301 SGQDQWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 QY 361 YTLVNDLDLVTAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 DB 361 YTLVNDLDLVTAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 5
 AAW95698
 ID AAW95698 standard; protein; 433 AA.

XX AAW95698;
 AC
 XX
 DT 16-JUN-1999 (first entry)
 XX
 DE Bacillus sp. Lion Y protease.
 XX
 XX PD498; subtilisin; Lion Y; protease; skin-care; feed; additive; soap;
 KW cosmetic; hair dye; sunscreens; acne; antiperspirants; insect repellent;
 KW deodorant; detergent; food; breadmaking; textile-treating; oral; dermal;
 KW pharmaceutical; agrochemical.
 XX
 XX Bacillus sp.
 OS
 XX WC9900489-A1.
 PN
 XX
 XX 07-JAN-1999.
 PD
 XX
 XX 22-JUN-1998; 98WO-DK000270.
 PF
 XX 25-JUN-1997; 97DK-00000753.
 PR
 XX 07-JUL-1997; 97US-0051830P.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX
 XX Olsen AA, Fatum TM, Deussen H, Roggen EL;
 PI WPI; 1999-095735/08.
 DR
 XX
 XX New modified polypeptide with attached low molecular weight polymer - has
 PT reduced respiratory allergenicity, useful in skin care products,
 PT detergents, as food additives or textile-treating compositions.
 PT
 XX
 XX Claim 10; Page 48-49; 60pp; English.
 PS
 XX
 XX The sequence is that of Lion Y protease. This can be used as an active
 CC ingredient: (i) in personal care products (especially skin-care products
 CC such as soaps, cosmetics, hair dyes, sunscreens, anti-acne products,
 CC antiperspirants, insect repellants or deodorants); (ii) in detergents (as
 CC laundry, dishwashing or hard-surface cleaners); (iii) food or feed
 CC additives (e.g. for breadmaking); (iv) in textile-treating compositions,
 CC or (v) in oral or dermal pharmaceuticals and agrochemicals
 CC
 XX SQ Sequence 433 AA;
 Query Match 99.3%; Score 2245; DB 2; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPFSGSLADPNHIAQFSSRGATRDGRIPKPDVTAPCTFTLSARSSLAPDSSFW 240
 DB 181 VGATENYRPFSGSLADPNHIAQFSSRGATRDGRIPKPDVTAPCTFTLSARSSLAPDSSFW 240
 QY 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVLGY 300
 DB 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVLGY 300
 QY 301 SGDCGWRVLDKSLNVAAYNEATATATGOKATYSFQAQAGKPKLSLWWTDPAGSTTAS 360
 DB 301 SGDCGWRVLDKSLNVAAYNEATATATGOKATYSFQAQAGKPKLSLWWTDPAGSTTAS 360

QY 361 YTLVNDLDELVTAPNGQKYVGNDFSPYDNNWNGNENVENFINAPQSGTYTIEVQAYNV 420
 DB 361 YTLVNDLDELVTAPNGQKYVGNDFSPYDNNWNGNENVENFINAPQSGTYTIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433
 RESULT 6
 AAY69207
 ID AAY69207 standard; protein; 433 AA.
 XX
 AC AAY69207;
 XX
 DT 30-MAY-2000 (first entry)
 XX
 DE Amino acid sequence of protease Lion Y.
 XX
 KW Protease Lion Y; polypeptide-polymer conjugate; washing performance;
 KW respiratory allergenicity; allergic reaction; detergent formulation;
 KW laundry; dishwashing; hard surface cleaner; agricultural chemical;
 KW skin care; cosmetic; oral pharmaceutical; dental pharmaceutical;
 KW textile processing.
 XX
 OS Bacillus sp.
 XX
 XX WO200004138-A1.
 PN
 XX 27-JAN-2000.
 PD
 XX
 XX 16-JUL-1999; 99WO-DK000406.
 PF
 XX
 XX 17-JUL-1998; 98DK-00000951.
 PR
 XX
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Bauditz P, Fatum TM, Olsen AA, Deussen H, Petersen DA;
 PI WPI; 2000-195024/17.
 DR
 XX
 XX New polypeptide-polymer conjugate, particularly enzyme conjugate, useful
 PT in detergent formulations.
 PT
 XX Disclosure; Page 74-76; 79pp; English.
 PS
 XX The present sequence represents a protease Lion Y protein. The protein
 CC may be used to produce the conjugates of the invention. The specification
 CC describes polypeptide-polymer conjugates which have improved washing
 CC performance and reduced respiratory allergenicity compared with the
 CC unconjugated polypeptide. The polymer provides, in water, a conformation
 CC that shields the molecular surface effectively, preventing association of
 CC antibodies that can induce an allergic reaction. The conjugates are used
 CC in industrial compositions, particularly detergent formulations (laundry,
 CC dishwashing or hard surface cleaners), but also in agricultural
 CC chemicals, skin care products (cosmetics and toiletries), oral and dental
 CC pharmaceuticals, or textile processing and treatment compositions
 CC
 XX SQ Sequence 433 AA;
 Query Match 99.3%; Score 2245; DB 3; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180

Db 121 GARIHNSWGPVNGAYTANSRQVDEYVRNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGLADNPNHIAQPSRGATDRGRIKPDVTAPGTTHLSARSLAPDSSFW 240
 Db 181 VGATENYRPSFGLADNPNHIAQPSRGATDRGRIKPDVTAPGTTHLSARSLAPDSSFW 240
 QY 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
 Db 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
 QY 301 SGDQGWGRVTLDKSLNVAYNVNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 Db 301 SGDQGWGRVTLDKSLNVAYNVNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 QY 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWDRNNVNFVINAPOSQTYTIEVOAYNV 420
 Db 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWDRNNVNFVINAPOSQTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIVH 433
 Db 421 PSGPQRFSLAIVH 433

RESULT 7

AA44619
 ID AAY44619 standard; protein; 433 AA.

XX AAY44619;
 XX 07-APR-2000 (first entry)

DT 07-APR-2000 (first entry)
 DE Bacillus Lion Y enzyme.

XX Lion Y enzyme; protease; allergic response; industrial composition;
 KW co-polymer; ethylene oxide; EO; propylene oxide; PO; conjugate;
 KW allergenicity; detergent; cosmetic; toiletries; textile treatment;
 KW agrochemical; pharmaceutical; food; feed additive.

XX Bacillus sp.
 XX WO9967370-A1.

XX 29-DEC-1999.
 XX 23-JUN-1999; 99WO-DK000359.

XX 23-JUN-1998; 98DK-00000809.
 XX (NOVO) NOVO-NORDISK AS.

XX Deussen H, Olsen AA, Fatum TM, Roggen EL;
 XX WPI; 2000-136981/12.

XX New conjugate of polypeptide, especially an enzyme, with copolymer of
 PT ethylene oxide and propylene oxide, used in e.g. cleaning compositions,
 PT has reduced allergenicity.
 XX Claim 10; Page 56-57; 62pp; English.

XX The present sequence is a Bacillus Lion Y enzyme, which is a protease
 CC capable of inducing an allergic response upon inhalation. The enzyme can
 CC be covalently coupled to a co-polymer comprising ethylene oxide (EO) and
 CC propylene oxide (PO) to reduce its allergenicity. This enzyme-polymer
 CC conjugate can be used in industrial compositions such as detergents,
 CC cosmetics, toiletries, textile treatment compositions, agrochemicals,
 CC oral and dermal pharmaceuticals and food and feed additives

XX SQ Sequence 433 AA;

Query Match 99.3%; Score 2245; DB 3; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;

Matches 423; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 Db 1 NDVARGIVKADVAQNNGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 QY 61 NNDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIWDSGGGLGGLPSNLTLFSQAWNA 120
 Db 61 NNDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIWDSGGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
 Db 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGLADNPNHIAQPSRGATDRGRIKPDVTAPGTTHLSARSLAPDSSFW 240
 Db 181 VGATENYRPSFGLADNPNHIAQPSRGATDRGRIKPDVTAPGTTHLSARSLAPDSSFW 240
 QY 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
 Db 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
 QY 301 SGDQGWGRVTLDKSLNVAYNVNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 Db 301 SGDQGWGRVTLDKSLNVAYNVNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
 QY 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWDRNNVNFVINAPOSQTYTIEVOAYNV 420
 Db 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWDRNNVNFVINAPOSQTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIVH 433
 Db 421 PSGPQRFSLAIVH 433

RESULT 8

AA50083
 ID AAM50083 standard; protein; 433 AA.

XX AAM50083;
 XX 12-AUG-2002 (first entry)

DT 12-AUG-2002 (first entry)
 DE Bacillus sp Y-(FERM BP-1029) alkaline protease protein fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX Bacillus sp.

XX EPI209233-A2.
 XX 29-MAY-2002.

XX 22-NOV-2001; 2001EP-00127851.
 XX 22-NOV-2000; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.
 XX (KAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX Okuda M, Saeki K;

XX WPI; 2002-437518/47.
 XX New modified alkaline proteases useful in detergent compositions.

XX Claim 5; Page 15-16; 25pp; English.
 XX This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

CC sequence represents a fragment of the alkaline protease Ya from *Bacillus*
 CC sp strain Y-(FERM BP-1029) described in the method of the invention

XX
 CC Sequence 433 AA;
 Query Match 99.1%; Score 2240; DB 5; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.7e-166;
 Matches 428; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNENGPNSGTISAPGTAKNAIT 180
 DB 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNENGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAPGTILSARSSLAPDSSFW 240
 DB 181 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAPGTILSARSSLAPDSSFW 240
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAAGATDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAAGATDVGLGYP 300
 QY 301 SGQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
 DB 301 NGDQGWGRVTLNKLNSVAVNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 420
 DB 361 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 420
 QY 421 PSQPQRFSLAIVH 433
 DB 421 PSQPQRFSLAIVH 433

RESULT 9
 AAW89548
 ID AAW89548 standard; protein; 636 AA.
 AC
 XX
 AC AAW89548;
 DT 12-APR-1999 (first entry)
 DE *Bacillus* sp. alkaline protease Y.
 KW Alkaline protease Y; detergent; surfactant; leather processing;
 XX debittering; flavour.
 OS *Bacillus* sp.
 XX
 DN WO9856927-A2.
 XX
 XX 17-DEC-1998.
 XX
 XX 09-JUN-1998; 98WO-US012005.
 XX
 XX 12-JUN-1997; 97US-00873479.
 PR
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 FA Sloma A, Christianson L;
 XX
 PI WPI; 1999-080908/07.
 XX
 XX Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing.

XX
 PS Claim 3; Page 55-56; 77pp; English.

XX
 CC This is the amino acid sequence of a *Bacillus* sp. alkaline protease Y
 CC that is said to have good alkali and surfactant resistance and improved
 CC detergency. It shows 77% identity to a newly isolated protease (see
 CC AAW89547) of *Bacillus* sp. JP170 (NCIB 12513). The invention provides
 CC vectors, recombinant host cells and methods for the recombinant
 CC production of such proteases. The protease are used in laundry and
 CC dishwashing detergents, for institutional and industrial cleaning, and
 CC for leather processing, as well as for debittering and enhancing the
 CC degree of hydrolysis of protein hydrolysates, for flavour development
 CC through hydrolysis of proteins, degradation of undesired peptides and in
 CC enzymatic synthesis of peptides. They have enhanced stability towards
 CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
 CC peroxy type. The invention also provides mutant cells in which the
 CC protease activity is diminished. Such cells can be used for the
 CC production of heterologous recombinant proteins

XX
 SQ Sequence 636 AA;

Query Match 98.9%; Score 2236; DB 2; Length 636;
 Best Local Similarity 98.8%; Pred. No. 5.9e-166;
 Matches 428; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 204 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 263
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 DB 264 NASDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 323
 QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNENGPNSGTISAPGTAKNAIT 180
 DB 324 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNENGPNSGTISAPGTAKNAIT 383
 QY 181 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAPGTILSARSSLAPDSSFW 240
 DB 384 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAPGTILSARSSLAPDSSFW 443
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAAGATDVGLGYP 300
 DB 444 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAAGATDVGLGYP 503
 QY 301 SGQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
 DB 504 SGQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 563
 QY 361 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 420
 DB 564 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 623
 QY 421 PSQPQRFSLAIVH 433
 DB 624 PSQPQRFSLAIVH 636

RESULT 10
 AAW50086
 ID AAW50086 standard; protein; 433 AA.
 AC
 XX
 AC AAW50086;
 DT 12-AUG-2002 (first entry)
 DE *Bacillus* sp. alkaline protease protein A-2 fragment.
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS *Bacillus* sp.
 XX
 XX EP1209233-A2.

XX 29-MAY-2002.
 XX 22-NOV-2001; 2001EP-00127851.
 XX 22-NOV-2000; 2000JP-00355165.
 PR 12-APR-2001; 2001JP-00114048.
 XX (KAOS) KAO CORP.
 XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX WPI; 2002-437518/47.
 XX New modified alkaline proteases useful in detergent compositions.
 PT Claim 5; Page 20-21; 25pp; English.
 XX This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease A-2 from Bacillus
 CC sp NCIB12513 described in the method of the invention
 XX NCIB12513 described in the method of the invention
 XX Sequence 433 AA;
 SQ
 Query Match 90.4%; Score 2044; DB 5; Length 433;
 Best Local Similarity 89.1%; Pred. No. 3.5e-151;
 Matches 386; Conservative 24; Mismatches 23; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSGGGLGGLPANLTLFQAWNA 120
 DB 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSGGGLGGLPANLTLFQAWNA 120
 QY 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKPDVTAPGTITLSARSSLAPDSSFW 240
 DB 181 VGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKPDVTAPGTITLSARSSLAPDSSFW 240
 QY 241 ANYNSKYAYNGGTSMATPIVAGNVAGLREHFVKNRGVTPKPSLLKAALIAGADYGLGFP 300
 DB 241 ANYNSKYAYNGGTSMATPIVAGNVAGLREHFVKNRGVTPKPSLLKAALIAGADYGLGFP 300
 QY 301 SGDQGWGRVTLDKSLNVAVYNEATATGKATYSFQAQKPLKISLVWTDAPGSTTAS 360
 DB 301 NGNQGWGRVTLDKSLNVAVFNETSPSTSQATYSFTAQAGKPLKISLVWSDAPGSTTAS 360
 QY 361 YTLVNDLDLVITAPNGKQVGNDFSPYDNNMDGRNNVENFINAPQSGTYTIEVQAVNV 420
 DB 361 LTLVNDLDLVITAPNGTKVGNDFAPYDNNMDGRNNVENFINAPQSGTYTIEVQAVNV 420
 QY 421 PSGPQRFSLAIVH 433
 DB 421 PVSPTQFSLAIVH 433
 RESULT 11
 AAW89547
 ID AAW89547 standard; protein; 641 AA.
 XX AAW89547;
 AC AAW89547;
 XX 12-APR-1999 (first entry)
 XX

DE Bacillus JPI70 protease.
 XX Protease; detergent; surfactant; leather processing; debittering;
 KW flavour.
 XX Bacillus sp.
 XX Key Location/Qualifiers
 PH Peptide 1..33
 FT /note= "signal peptide"
 FT Region 34..208
 FT /note= "prepro region"
 FT Protein 209..641
 FT /note= "mature protein"
 XX WO9856927-A2.
 PN 17-DEC-1998.
 PD 09-JUN-1998; 98WO-US012005.
 XX 12-JUN-1997; 97US-00873479.
 PR (NOVO) NOVO NORDISK BIOTECH INC.
 PA Sloma A, Christianson L;
 PI WPI; 1999-080908/07.
 DR N-PSDB; AAV82382.
 DR Novel protease from Bacillus subtilis LC20 - useful in laundry and
 XX dishwashing detergents and for leather processing.
 XX Claim 7; Page 53-54; 77pp; English.
 XX This is the amino acid sequence of a novel protease of Bacillus sp. JPI70
 CC (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene
 CC (see AAV82382). The entire protein, including the signal peptide and
 CC prepro region, has 77% identity to alkaline protease Y (see AAW89548)
 CC from Bacillus. The invention provides vectors, recombinant host cells and
 CC methods for the recombinant production of the protease. The protease is
 CC used in laundry and dishwashing detergents, for institutional and
 CC industrial cleaning, and for leather processing, as well as for
 CC debittering and enhancing the degree of hydrolysis of protein
 CC hydrolysates, for flavour development through hydrolysis of proteins.
 CC degradation of undesired peptides and in enzymatic synthesis of peptides.
 CC It has enhanced stability towards oxidation under alkaline conditions,
 CC e.g. towards bleaching agents of the peroxy type. The invention also
 CC provides mutant cells in which the protease activity is diminished. Such
 CC cells can be used for the production of heterologous recombinant proteins
 XX Sequence 641 AA;
 SQ
 Query Match 90.4%; Score 2044; DB 2; Length 641;
 Best Local Similarity 89.1%; Pred. No. 6.1e-151;
 Matches 386; Conservative 24; Mismatches 23; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 209 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 268
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSGGGLGGLPANLTLFQAWNA 120
 DB 269 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSGGGLGGLPANLTLFQAWNA 328
 QY 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 329 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 388
 QY 181 VGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKPDVTAPGTITLSARSSLAPDSSFW 240
 DB 389 VGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKPDVTAPGTITLSARSSLAPDSSFW 448

QY 241 ANYNSKYAMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLKAALIAAGTAVGLGYP 300
 Db 449 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLKAALIAAGTAVGLGYP 508
 QY 301 SGOQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 360
 Db 509 NGNQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 568
 QY 361 YTLVNDLVLITAPNGQKYGNDPFPYDNNWGRNENVENFINAPSGTITVEVQAYN 420
 Db 569 LTLVNDLVLITAPNGQKYGNDPFPYDNNWGRNENVENFINAPSGTITVEVQAYN 628
 QY 421 PPSQPFSLAIHV 433
 Db 629 PVSPQPFSLAIHV 641

RESULT 12
 AAM50085
 ID AAM50085 standard; protein; 434 AA.
 XX
 AC AAM50085;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Bacillus sp alkaline protease protein A-1 fragment.
 XX
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS Bacillus sp.
 XX
 PN EP1209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-00127851.
 XX
 PR 22-NOV-2000; 2000JP-00355166.
 PR 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.
 XX
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 5; Page 18-19; 25pp; English.
 XX
 CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease A-1 from Bacillus
 CC sp NCIB12289 described in the method of the invention
 XX
 SQ Sequence 434 AA;
 Query Match 89.5%; Score 2024.5; DB 5; Length 434;
 Best Local Similarity 88.5%; Pred. No. 1.2e-149;
 Matches 384; Conservative 28; Mismatches 21; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLTGRNDSMHEAFRGKITALYALGRN 60
 Db 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLTGRNDSMHEAFRGKITALYALGRN 60
 QY 61 NANDPNHGHTHAGSVLGNAL-NKGWAPQANLVFQSIMDSSGGLGGLPNSNLTFLSQAWN 119
 Db 61 NANDPNHGHTHAGSVLGNAL-NKGWAPQANLVFQSIMDSSGGLGGLPNSNLTFLSQAYS 120
 QY 120 AGARIHTNSWGAPVNGAYTANSRQVDYFVRNMDTVLFAACNENGPNSGTISAPGTAKNAI 179

Db 121 AGARIHTNSWGAPVNGAYTANSRQVDYFVRNMDTVLFAACNENGPNSGTISAPGTAKNAI 180
 QY 180 TVGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLPDSSF 239
 Db 181 TVGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLPDSSF 240
 QY 240 WANYNSKYAMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLKAALIAAGTAVGLGYP 299
 Db 241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLKAALIAAGTAVGLGYP 300
 QY 300 PSCDQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 359
 Db 301 PSCDQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 360
 QY 360 SVTLVNDLVLITAPNGQKYGNDPFPYDNNWGRNENVENFINAPSGTITVEVQAYN 419
 Db 361 SVTLVNDLVLITAPNGQKYGNDPFPYDNNWGRNENVENFINAPSGTITVEVQAYN 420
 QY 420 VPSQPFSLAIHV 433
 Db 421 VPSQPFSLAIHV 434

RESULT 13
 AAM50081
 ID AAM50081 standard; protein; 434 AA.
 XX
 AC AAM50081;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Bacillus sp KSM-KP9860 alkaline protease protein fragment.
 XX
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS Bacillus sp.
 XX
 PN EP1209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-00127851.
 XX
 PR 22-NOV-2000; 2000JP-00355166.
 PR 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.
 XX
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 5; Page 12-13; 25pp; English.
 XX
 CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC Bacillus sp strain KSM-KP9860 described in the method of the invention
 XX
 SQ Sequence 434 AA;
 Query Match 88.9%; Score 2010.5; DB 5; Length 434;
 Best Local Similarity 88.5%; Pred. No. 1.5e-148;
 Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLTGRNDSMHEAFRGKITALYALGRN 60

Db 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNGHGTTHVAGSVLGN-ALNKGWAPQANLVFOSIMDSSGGLGGLPSNLNTLFSQAWN 119
 Db 61 NANDTNGHGTTHVAGSVLGN-ALNKGWAPQANLVFOSIMDSSGGLGGLPSNLNTLFSQAFS 120
 QY 120 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLFAAGNEGPNSTISAPGTAKNAI 179
 Db 121 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLFAAGNEGPNSTISAPGTAKNAI 180
 QY 180 TVGATENYRPSFGSLADNPNHIAQFSSRGATDGRKPDVAPGTFILSARSSSLAPDSSF 239
 Db 181 TVGATENLRFPSFGSYADNINHVAQFSSRGFTKDGRIKPDVAPGTFILSARSSSLAPDSSF 240
 QY 240 WANYNSKYAYMGGTSMATPIVAGNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 299
 Db 241 WANHDSKYAYMGGTSMATPIVAGNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 300
 QY 300 PSGDQGWGRVTLDKSLNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 359
 Db 301 PNGNQGWGRVTLDKSLNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 360
 QY 360 SVTLVNDLVLITAPNGQYKVGNDYFSDYDNNWGRNNDVFNAPQSGTYTIEVQAYN 419
 Db 361 SVTLVNDLVLITAPNGQYKVGNDYFSDYDNNWGRNNDVFNAPQSGTYTIEVQAYN 420
 QY 420 VPSGPQRFSLAIVH 433
 Db 421 VPVGPQNFSLAIVN 434

RESULT 14

AAV17089
 ID AAV17089 standard; protein; 639 AA.

XX AAV17089;

AC AAV17089;

DT 20-MAR-2003 (revised)

DT 21-JUL-1999 (first entry)

XX Bacillus alkaline protease.

XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;

KW washing composition; oxidising agent.

XX Bacillus sp.

OS Bacillus sp.

XX WO9918218-A1.

XX 15-APR-1999.

XX 07-OCT-1998; 98WO-JP004528.

XX 07-OCT-1997; 97JP-00274570.

XX (KAOS) KAO CORP.

XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;

PI Shikata S, Nomura M;

XX WPI; 1999-287736/27.

DR N-PSDB; AAX37277.

XX Alkali protease from Bacillus used in washing powders.

PT Disclosure; Page 53-58; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of

XX Bacillus. The proteases ability to digest casein is not inhibited by

XX oleic acid and they have a high stability to oxidising agents. The

XX alkaline protease of the invention has the following properties: (a) it

XX is active over the pH range 4-13 and has at least 80% of its optimum

XX activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is

CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease. (Updated
 CC on 20-MAR-2003 to correct DR field.)

SQ Sequence 639 AA;

Query Match 88.6%; Score 2002.5; DB 2; Length 639;

Best Local Similarity 88.2%; Pred. No. 1.1e-147;

Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60

Db 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 265

QY 61 NANDPNGHGTTHVAGSVLGN-ALNKGWAPQANLVFOSIMDSSGGLGGLPSNLNTLFSQAWN 119

Db 266 NANDTNGHGTTHVAGSVLGN-ALNKGWAPQANLVFOSIMDSSGGLGGLPSNLNTLFSQAFS 325

QY 120 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLFAAGNEGPNSTISAPGTAKNAI 179

Db 326 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLFAAGNEGPNSTISAPGTAKNAI 385

QY 180 TVGATENYRPSFGSLADNPNHIAQFSSRGATDGRKPDVAPGTFILSARSSSLAPDSSF 239

Db 386 TVGATENLRFPSFGSYADNINHVAQFSSRGFTKDGRIKPDVAPGTFILSARSSSLAPDSSF 445

QY 240 WANYNSKYAYMGGTSMATPIVAGNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 299

Db 446 WANHDSKYAYMGGTSMATPIVAGNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 505

QY 300 PSGDQGWGRVTLDKSLNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 359

Db 506 PNGNQGWGRVTLDKSLNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 565

QY 360 SVTLVNDLVLITAPNGQYKVGNDYFSDYDNNWGRNNDVFNAPQSGTYTIEVQAYN 419

Db 566 SVTLVNDLVLITAPNGQYKVGNDYFSDYDNNWGRNNDVFNAPQSGTYTIEVQAYN 625

QY 420 VPSGPQRFSLAIVH 433

Db 626 VPVGPQNFSLAIVN 639

RESULT 15

AAV17091
 ID AAV17091 standard; protein; 640 AA.

XX AAV17091;

AC AAV17091;

DT 20-MAR-2003 (revised)

DT 21-JUL-1999 (first entry)

XX Bacillus alkaline protease.

XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;

KW washing composition; oxidising agent.

XX Bacillus sp.

OS Bacillus sp.

XX WO9918218-A1.

XX 15-APR-1999.

XX 07-OCT-1998; 98WO-JP004528.

XX 07-OCT-1997; 97JP-00274570.

XX (KAOS) KAO CORP.

XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX WPI; 1999-287736/27.
DR N-PSDB; AAX37279.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Disclosure; Page 63-68; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX
XX SQ Sequence 640 AA;

Query Match 88.4%; Score 1999.5; DB 2; Length 640;
Best Local Similarity 87.8%; Pred. No. 1.8e-147;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQNNYGLYQGQWVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
207 NDVARGIVKADVQSSYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
QY 61 NANDPNHGHTHAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGLGLPSNLTFLSQAWN 119
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
267 NANDTNGHGTAVAGSVLGNSTNKGMAPQANLVFQSIMDSSGGLGLPSNLTFLSQAYS 326
QY 120 AGARIHTNSGAPVNGAYTANSQVDEYVNRNDMTVLFAAGNEGPNSTISAPGTAKNAI 179
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
327 AGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI 386
QY 180 TVGATENYRPSFGSLADPNHIAQFSSRGATRDGRKPKDVTAPGTFTLSARSSLAPDSSF 239
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
387 TVGATENLRPSFGYADNINHVAFSSRGPTKDGRIKPDVWAPGTFTLSARSSLAPDSSF 446
QY 240 WANTNSKYANGGTSMTPIVAGNVAQLREHFVKRGITPKPSLIKAALIAGATDVGLGY 299
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
447 WANHDSKYANGGTSMTPIVAGNVAQLREHFVKRGITPKPSLIKAALIAGADIGLGY 506
QY 300 PSGDQGWGRVTLDXSLNVAVYNEATATGOKATYSFQAQAGKPKISLVMTDAPGSTTA 359
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
507 PNGQGWGRVTLDXSLNVAVYNESSLSLTSQATYSFTATAGKELKISLVWSDAPASTTA 566
QY 360 SYTLVNDLVLITAPNGOKYVGNDFSPYDNNWGRNNVNFINAPQSGTYTIEVQAYN 419
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
567 SVTLVNDLVLITAPNGTYVGNDFSPYDNNWGRNNVNFINAPQSGTYTIEVQAYN 626
QY 420 VPSGQFQFSLAIVH 433
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
627 VPGQFQFSLAIVN 640

Search completed: March 31, 2004, 16:04:32
Job time : 48.1304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:59:39 ; Search time 13.9892 Seconds
(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNYGLY.....EVAQNVPSGQRSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/iaa/6CTUS-COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2245	99.3	433	4	US-09-104-623A-4
2	2245	99.3	433	4	US-09-019-532-4
3	2245	99.3	433	4	US-09-338-746-4
4	2245	99.3	635	2	US-08-873-479-43
5	2044	90.4	641	2	US-08-873-479-42
6	2002.5	88.6	639	4	US-09-509-814A-4
7	1999.5	88.4	640	4	US-09-509-814A-8
8	1998.5	88.4	640	4	US-09-509-814A-6
9	1952.5	86.4	639	4	US-09-509-814A-1
10	1952.5	86.4	640	4	US-09-509-814A-2
11	1514	67.0	345	4	US-08-512-251A-10
12	1514	67.0	345	4	US-09-515-150A-10
13	1514	67.0	345	4	US-09-196-281-13
14	443	19.6	659	3	US-08-894-818B-1
15	443	19.6	659	4	US-09-445-472-12
16	422.5	18.7	522	4	US-09-445-472-1
17	422.5	18.7	522	3	US-08-894-818B-3
18	422.5	18.7	522	4	US-09-445-472-4
19	422.5	18.7	654	3	US-08-894-818B-35
20	422.5	18.7	654	4	US-09-445-472-16
21	401.5	17.8	659	3	US-08-894-818B-5
22	343	15.2	520	4	US-09-000-016-7
23	343	15.2	520	4	US-09-514-340-7
24	343	15.2	734	3	US-09-000-016-4
25	343	15.2	734	4	US-08-514-340-4
26	343	15.2	823	3	US-09-000-016-2
27	343	15.2	823	4	US-09-514-340-2

28	307	13.6	903	1	US-08-750-532-1	Sequence 1, Appli
29	307	13.6	1398	1	US-08-750-532-9	Sequence 2, Appli
30	307	13.6	1398	3	US-08-894-818B-8	Sequence 8, Appli
31	307	13.6	1398	4	US-09-445-472-6	Sequence 6, Appli
32	280.5	12.4	237	1	US-08-750-532-18	Sequence 18, Appli
33	272	12.0	275	2	US-08-750-406A-1	Sequence 1, Appli
34	272	12.0	275	3	US-09-327-118-1	Sequence 1, Appli
35	269	11.9	269	1	US-08-431-387-5	Sequence 5, Appli
36	269	11.9	269	1	US-08-431-387-6	Sequence 6, Appli
37	269	11.9	269	1	US-08-322-677A-10	Sequence 10, Appli
38	269	11.9	269	1	US-08-322-676-10	Sequence 10, Appli
39	269	11.9	269	2	US-08-140-083A-10	Sequence 10, Appli
40	269	11.9	269	3	US-08-898-218-10	Sequence 10, Appli
41	269	11.9	269	3	US-08-848-793-10	Sequence 10, Appli
42	269	11.9	269	3	US-09-255-502-5	Sequence 5, Appli
43	269	11.9	269	3	US-09-024-532-3	Sequence 3, Appli
44	269	11.9	269	3	US-08-369-050-4	Sequence 4, Appli
45	269	11.9	269	3	US-08-090-207-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-104-623A-4
; Sequence 4, Application US/09104623A
; Patent No. 6303752
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Fatum, Tine Muxoll
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: A Modified Polypeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 63037520 No. 6303752disk of No. 6303752th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,623A
; FILING DATE: 25-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/POCKET NUMBER: 5256.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-857-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: Bacillus sp. Y
; US-09-104-623A-4

Query Match 99.3%; Score 2245; DB 4; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.6e-166;
Matches 429; Conservative 1; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNYGLYGGQVAVADTGLDTGRNDSHHEAFRGKITALVALGRTN 60

```
Db 1 NDVARGIVKADVAQNNGYGLYGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTHTVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSRQVDYVYRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db 121 GARIHTNSGAPVNGAYTANSRQVDYVYRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPKDVTAPOGTFTLSARSSSLAPDSSEW 240
Db 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPKDVTAPOGTFTLSARSSSLAPDSSEW 240
QY 241 ANYSKYAYMGTSMTATPIVAGNVAQLREHFIKNRGITPKPSLIKAALIAGATDVGLGYP 300
Db 241 ANYSKYAYMGTSMTATPIVAGNVAQLREHFIKNRGITPKPSLIKAALIAGATDVGLGYP 300
QY 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 360
Db 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 360
QY 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNNVENVFNAPOSQGTYYIEVQAYNV 420
Db 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNNVENVFNAPOSQGTYYIEVQAYNV 420
QY 421 PSGQRFSLAIVH 433
Db 421 PSGQRFSLAIVH 433
```

```
RESULT 2
US-09-019-532-4
; Sequence 4, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Olsen, Annette
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922.204-US
; CURRENT APPLICATION NUMBER: US/09/019,532B
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0038/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-019-532-4
```

```
Query Match 99.3%; Score 2245; DB 4; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.6e-166;
Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYGLYGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYGLYGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTHTVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSRQVDYVYRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
```

```
Db 121 GARIHTNSGAPVNGAYTANSRQVDYVYRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPKDVTAPOGTFTLSARSSSLAPDSSEW 240
Db 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPKDVTAPOGTFTLSARSSSLAPDSSEW 240
QY 241 ANYSKYAYMGTSMTATPIVAGNVAQLREHFIKNRGITPKPSLIKAALIAGATDVGLGYP 300
Db 241 ANYSKYAYMGTSMTATPIVAGNVAQLREHFIKNRGITPKPSLIKAALIAGATDVGLGYP 300
QY 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 360
Db 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 360
QY 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNNVENVFNAPOSQGTYYIEVQAYNV 420
Db 361 YTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNNVENVFNAPOSQGTYYIEVQAYNV 420
QY 421 PSGQRFSLAIVH 433
Db 421 PSGQRFSLAIVH 433
```

```
RESULT 3
US-09-338-746-4
; Sequence 4, Application US/09338746
; Patent No. 6638526
; GENERAL INFORMATION:
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Olsen, Arne A.
; APPLICANT: Fatum, Tine M.
; APPLICANT: Roggen, Erwin L.
; TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
; FILE REFERENCE: 5619.200-US
; CURRENT APPLICATION NUMBER: US/09/338,746
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: PA 1998 00809
; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 60/091,461
; EARLIER FILING DATE: 1998-07-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-338-746-4
```

```
Query Match 99.3%; Score 2245; DB 4; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.6e-166;
Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYGLYGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYGLYGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTHTVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSRQVDYVYRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db 121 GARIHTNSGAPVNGAYTANSRQVDYVYRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPKDVTAPOGTFTLSARSSSLAPDSSEW 240
Db 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPKDVTAPOGTFTLSARSSSLAPDSSEW 240
QY 241 ANYSKYAYMGTSMTATPIVAGNVAQLREHFIKNRGITPKPSLIKAALIAGATDVGLGYP 300
Db 241 ANYSKYAYMGTSMTATPIVAGNVAQLREHFIKNRGITPKPSLIKAALIAGATDVGLGYP 300
QY 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVMTDAPGSTTAS 360
```

Db 301 SDQGWGRVTLDKSLNVAIVNEATATATGQKATYSFQAQKPKLISLVWTDAPGSTAS 360
QY 361 YTLVNDLDLVTAPNGQKYVGNDFSPYDNNWGDGRNNVNFVINAPOSQGTITIEVQAYNV 420
Db 361 YTLVNDLDLVTAPNGQKYVGNDFSPYDNNWGDGRNNVNFVINAPOSQGTITIEVQAYNV 420
QY 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

RESULT 4
US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-873-479-43

Query Match 99.3%; Score 2245; DB 2; Length 635;
Best Local Similarity 99.1%; Pred. No. 2.7e-166;
Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAAQNNGYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 203 NDVARGIVKADVAAQNNGYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 262

QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSSGGLGGLPSNLTLFSQAWNA 120
Db 263 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSSGGLGGLPSNLTLFSQAWNA 322

QY 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db 323 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 382

QY 181 VGATENYRPSFGSIADPNHIAQFSRGATRDGRIPKPDVTAPCTFTLSARSSILAPDSSFW 240
Db 383 VGATENYRPSFGSIADPNHIAQFSRGATRDGRIPKPDVTAPCTFTLSARSSILAPDSSFW 442

QY 241 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFITKRGITPKPSLIKAALIAGATDVLGLYP 300
Db 443 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFITKRGITPKPSLIKAALIAGATDVLGLYP 502
QY 301 SGDOGWGRVTLDKSLNVAIVNEATATATGQKATYSFQAQKPKLISLVWTDAPGSTAS 360
Db 503 SGDOGWGRVTLDKSLNVAIVNEATATATGQKATYSFQAQKPKLISLVWTDAPGSTAS 562
QY 361 YTLVNDLDLVTAPNGQKYVGNDFSPYDNNWGDGRNNVNFVINAPOSQGTITIEVQAYNV 420
Db 563 YTLVNDLDLVTAPNGQKYVGNDFSPYDNNWGDGRNNVNFVINAPOSQGTITIEVQAYNV 622
QY 421 PSGPQRFSLAIVH 433
Db 623 PSGPQRFSLAIVH 635

RESULT 5
US-08-873-479-42
; Sequence 42, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-873-479-42

Query Match 90.4%; Score 2044; DB 2; Length 641;
Best Local Similarity 89.1%; Pred. No. 1.1e-150;
Matches 386; Conservative 24; Mismatches 23; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAAQNNGYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 209 NDVARGIVKADVAAQNNGYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 268

QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSSGGLGGLPSNLTLFSQAWNA 120
Db 269 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSSGGLGGLPSNLTLFSQAWNA 328

QY 121 GARIHTNSGAPVNGAYTANSQVDEYVENNDMTVLEAGNEGNSGTISAPGTAKNAIT 180
Db 329 GARIHTNSGAPVNGAYTANSQVDEYVENNDMTVLEAGNEGNSGTISAPGTAKNAIT 388
QY 181 VGATENRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSF 240
Db 389 VGATENRPSFGSYADNINHVAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSF 448
QY 241 ANYNSKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 300
Db 449 AHDHDKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 508
QY 301 SGOQGWGRVTLDKSLNVAYNEATATGQKATYSFOAQGKPLKISLVWTDAPGSTTA 360
Db 509 NGNOGWGRVTLDKSLNVAFVNETSPLSTSQKATYSFOAQGKPLKISLVWTDAPGSTTA 568
QY 361 YTLVNDLDELITAPNGQKVGNDPSYPYDNNWGRNENVENFINAPOSGTITIEVQAYN 420
Db 569 LTLVNDLDELITAPNGTKYVNDFTAPYDNNWGRNENVENFINAPOSGTITIEVQAYN 628
QY 421 PSQPORFSLAIVH 433
Db 629 PVSQPTFSLAIVH 641

RESULT 6

US-09-509-814A-4
; Sequence 4, Application US/09509814A
; Patent No. 6376227

GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR FILING DATE: 1998-10-07
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.

US-09-509-814A-4

Query Match 88.6%; Score 2002.5; DB 4; Length 639;
Best Local Similarity 88.2%; Pred. No. 1.9e-147;
Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 206 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 265
QY 61 NADPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSGGGLGLPSNLTLFQAWN 119
Db 266 NADPNHGHTHVAGSVLGNATNKGMAPQANLVFQSIMDSGGGLGLPSNLTLFQAFS 325
QY 120 AGARIHTNSWCAVNGAYTANSROVDYVRNNDMTVLFAGNEGPNSGTISAPGTAKNAI 179
Db 326 AGARIHTNSWCAVNGAYTDSRVDDYVRKNDMTILFAAGNEPNNGGTISAPGTAKNAI 385
QY 180 TVGATENRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSF 239

Db 386 TVGATENRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAPGTIYLSARSLAPDSF 445
QY 240 WANYNSKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 299
Db 446 WANDHDKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGADVGLGY 505
QY 300 PSQDQGWGRVTLDKSLNVAYNEATATGQKATYSFOAQGKPLKISLVWTDAPGSTTA 359
Db 506 PNGNQGWGRVTLDKSLNVAYNESSALSTSQKATYFTTAGKPLKISLVWSDAPASTTA 565
QY 360 SYTLVNDLDELITAPNGQKVGNDPSYPYDNNWGRNENVENFINAPOSGTITIEVQAYN 419
Db 566 SVTLVNDLDELITAPNGTRYVGNDFPAPDNNWGRNENVENFINPSQSGTITIEVQAYN 625
QY 420 VPSQPORFSLAIVH 433
Db 626 VPVGPQNFSLAIVN 639

RESULT 7

US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227

GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR FILING DATE: 1998-10-07
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.

US-09-509-814A-8

Query Match 88.4%; Score 1999.5; DB 4; Length 640;
Best Local Similarity 87.8%; Pred. No. 3.2e-147;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
QY 61 NADPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSGGGLGLPSNLTLFQAWN 119
Db 267 NADPNHGHTHVAGSVLGNSTNKGMAPQANLVFQSIMDSGGGLGLPSNLTLFQAYS 326
QY 120 AGARIHTNSWCAVNGAYTANSROVDYVRNNDMTVLFAGNEGPNSGTISAPGTAKNAI 179
Db 327 AGARIHTNSWCAVNGAYTDSRVDDYVRKNDMTILFAAGNEPNNGGTISAPGTAKNAI 386
QY 180 TVGATENRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSF 239
Db 387 TVGATENRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVTPAGTIFLSARSLAPDSF 446
QY 240 WANYNSKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 299
Db 447 WANDHDKYAVMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGADVGLGY 506
QY 300 PSQDQGWGRVTLDKSLNVAYNEATATGQKATYSFOAQGKPLKISLVWTDAPGSTTA 359

```
Db 507 PNEGQGWGRVLDLSDSLNVAAYNNESSLSSTQKATVSFTATAGKPLKISLVWSDAPASTTA 566
Qy 360 SYTLVNDLDLVTAPNGOKYVGNDFSPYDNNWGRNNVENVFINAPQSGTYTIEVQAYN 419
Db 567 SVTLVNDLDLVTAPNGTYVGNDFSPYDNNWGRNNVENVFINAPQSGTYTIEVQAYN 626
Qy 420 VPSGPQSFSLAIVH 433
Db 627 VPVGPQSFSLAIVN 640

RESULT 8
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, MASAFUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 840
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6

Query Match 88.4%; Score 1998.5; DB 4; Length 640;
Best Local Similarity 87.8%; Pred. No. 3 8e-147;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQNNGYLGQSVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 60
Db 207 NDVARGIVKADVAQSSVGLYQSQIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 266
Qy 61 NANDPNHGHTHVAGSVLGN-ALNKGWAPQANLVFQSIMDSSGGLGSLPSNLNTLFSQAWN 119
Db 267 NANDTNGHGHVAGSVLGNSTNGKWAQANLVFQSIMDSSGGLGSLPSNLNTLFSQAYS 326
Qy 120 AGARIHNTSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAI 179
Db 327 AGARIHNTSGAAVNGAYTTDSRVDDYVRKNDMTVLFAAGNEGPNSTISAPGTAKNAI 386
Qy 180 TVGATENRPSFGSLADNPNHIAQFSRGATRCRKPDPVAPGTIFLSARSLAPDSF 239
Db 387 TVGATENRPSFGSLADNPNHIAQFSRGATRCRKPDPVAPGTIFLSARSLAPDSF 446
Qy 240 WANVNSKYAVNGTSMATPTVAGNVAQLREHFKRNGITPKPSLIKAALIAGATDVGLGY 299
Db 447 WANHDSKYAVNGTSMATPTVAGNVAQLREHFKRNGITPKPSLLKAALIAGAADIGLY 506
Qy 300 PSGDQGWGRVTLDKSLNVAVNBATALATQCKATYSFQAQAGKPLKISLVWTDAPGSTTA 359
Db 507 PNEGQGWGRVTLDKSLNVAVNBESSLSSTQKATVSFTATAGKPLKISLVWSDAPASTTA 566
Qy 360 SYTLVNDLDLVTAPNGOKYVGNDFSPYDNNWGRNNVENVFINAPQSGTYTIEVQAYN 419
Db 567 SVTLVNDLDLVTAPNGTYVGNDFSPYDNNWGRNNVENVFINAPQSGTYTIEVQAYN 626
```

```
Qy 420 VPSGPQSFSLAIVH 433
Db 627 VPVGPQSFSLAIVN 640

RESULT 9
US-09-509-814A-1
; Sequence 1, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (46)..(46)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (105)..(105)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (130)..(130)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
```

LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 86.4%; Score 1952.5; DB 4; Length 639;
Best Local Similarity 86.6%; Pred. No. 1.4e-143;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGTGRNDSSMHEAFRGKITALYALGRTN 60
Db 206 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGTGRNDSSMHEAFRGKITALYALGRTN 265
Qy 61 NANDPNHGHTHVAGSVLGN-ALNKGWAPQANLVFQSIMDSSGGGLGPLSNLTLFSQAWN 119
Db 266 NANDTNGHTHVAGSVLGNXTNKGWAPQANLVFQSIMDSSGGGLGPLSNLTLFSQAXS 325
Qy 120 AGARIHTNSGAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNSGTISAFGTAKNAI 179
Db 326 AGARIHTNSGAAVNGAYTTDSRVDYVRKNMTVLFAAGNEXPNGGTISAFGTAKNAI 385
Qy 180 TVGATENYRPSFGSLADNPNHIAQFSRGATROGRIPKDVTAFTGTFILSARSLADPSSF 239
Db 386 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTXILSARSLADPSSF 445
Qy 240 WANYNSKYAYMGTSMATPIVAGNVAQLREHFKNRGITPKPSLIKAALITAGATDVLGY 299
Db 446 WANHDSKIAYMGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALITAGADLGY 505
Qy 300 PSGDQGWGRVTLDKSLNVAYVNEATALATQCKATYSFQAQAGKPLKISLVWTDAPGSTTA 359
Db 506 PNGNQGWGRVTLDKSLNVAYVNESSXLSTSQATYATATAGKPLKISLVWDAPASTTA 565
Qy 360 SYTLVNDLDELVITAPNGQKYVGNDFSPYDNNWDGRNNVENVTINAPQSGCTYIEVOAYN 419
Db 566 SYTLVNDLDELVITAPNGTXYVGNDFXXPXXXNWDGRNNVENFINPQSGCTYIEVOAYN 625
Qy 420 VPSGQRFSLAIVH 433
Db 626 VPVGQXFSLAIVN 639

RESULT 10
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08

```

; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (30)..(30)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (33)..(33)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (48)..(48)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (54)..(54)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (71)..(71)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (75)..(75)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (90)..(90)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (103)..(103)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (106)..(106)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (129)..(129)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (131)..(131)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (132)..(132)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (133)..(133)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (134)..(134)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (147)..(147)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (149)..(149)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (161)..(161)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (166)..(166)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (173)..(173)
; OTHER INFORMATION: Xaa is any amino acid

```

```

; NAME/KEY: misc_feature
; LOCATION: (184)..(184)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (189)..(189)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (190)..(190)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (195)..(195)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (287)..(287)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (325)..(325)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (370)..(370)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (432)..(432)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (502)..(502)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (532)..(532)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (542)..(542)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (585)..(585)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (592)..(592)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (593)..(593)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (595)..(595)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (596)..(596)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (597)..(597)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (612)..(612)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (633)..(633)
; OTHER INFORMATION: Xaa is any amino acid
;
; US-09-509-814A-2

```

Query Match 86.4%; Score 1952.5; DB 4; Length 640;
Best Local Similarity 86.6%; Pred. No. 1.4e-143;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQNNGLYGCGGVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 60
DB 207 NDVARGIVKADVAQNNGLYGCGGVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 266

QY 61 NNDPNGHGHVAGSVLGN -ALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 119
Db 267 NNDPNGHGHVAGSVLGN -ALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 326
QY 120 AGARHTNSWGAPVNGAYTANSROVDYVRNNDMTLFAAGNEGPNSTISAPGTAKNAI 179
Db 327 AGARHTNSWGAPVNGAYTANSROVDYVRNNDMTLFAAGNEGPNSTISAPGTAKNAI 386
QY 180 TVGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 239
Db 387 TVGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 446
QY 240 WANYNSKYAYMGTSMTATPIVAGNVAOLREHFVKNRGVTPKPSLTKAALIAGATDVGLGYP 299
Db 447 WANYNSKYAYMGTSMTATPIVAGNVAOLREHFVKNRGVTPKPSLTKAALIAGATDVGLGYP 506
QY 300 PPSGQGWGRVTLDKSLNVAAYNEATATGQKATYSFQAQAGKPLKISLWMTDAPGSTTA 359
Db 507 PPSGQGWGRVTLDKSLNVAAYNEATATGQKATYSFQAQAGKPLKISLWMTDAPGSTTA 566
QY 360 SVTLVNDLVLITAPNGQKQVGNDRSPYDNNWDGNNVNFVNAPOSGYTIIEVQAYN 419
Db 567 SVTLVNDLVLITAPNGQKQVGNDRSPYDNNWDGNNVNFVNAPOSGYTIIEVQAYN 626
QY 420 VPSGQGWGRVTLDKSLNVAAYNEATATGQKATYSFQAQAGKPLKISLWMTDAPGSTTA 433
Db 627 VPSGQGWGRVTLDKSLNVAAYNEATATGQKATYSFQAQAGKPLKISLWMTDAPGSTTA 640

RESULT 11
US-09-512-251A-10
; Sequence 10, Application US/09512251A
; Patent No. 655335

GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349.204-US
; CURRENT APPLICATION NUMBER: US/09/512,251A
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1

SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-512-251A-10

Query Match 67.0%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 7e-110;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 29 NDVARGIVKADVAQNNGYLGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 88
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 120
Db 89 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 148
QY 121 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTLFAAGNEGPNSTISAPGTAKNAIT 180
Db 149 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTLFAAGNEGPNSTISAPGTAKNAIT 208
QY 181 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 240
Db 209 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 268
QY 241 ANYNSKYAYMGTSMTATPIVAGNVAOLREHFVKNRGVTPKPSLTKAALIAGATDVGLGYP 300
Db 269 ANYNSKYAYMGTSMTATPIVAGNVAOLREHFVKNRGVTPKPSLTKAALIAGATDVGLGYP 328

QY 301 SGDQGWGRVTLDKSLNV 317
Db 329 NGNQGWGRVTLDKSLNV 345

RESULT 12

US-09-515-150A-10
; Sequence 10, Application US/09515150A
; Patent No. 6558938
GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5348.204-US
; CURRENT APPLICATION NUMBER: US/09/515,150A
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-515-150A-10

Query Match 67.0%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 7e-110;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 29 NDVARGIVKADVAQNNGYLGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 88
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 120
Db 89 NNDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 148
QY 121 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTLFAAGNEGPNSTISAPGTAKNAIT 180
Db 149 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTLFAAGNEGPNSTISAPGTAKNAIT 208
QY 181 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 240
Db 209 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 268
QY 241 ANYNSKYAYMGTSMTATPIVAGNVAOLREHFVKNRGVTPKPSLTKAALIAGATDVGLGYP 300
Db 269 ANYNSKYAYMGTSMTATPIVAGNVAOLREHFVKNRGVTPKPSLTKAALIAGATDVGLGYP 328

RESULT 13

US-09-196-281-13
; Sequence 13, Application US/09196281A
; Patent No. 6605458
GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/09/196,281A
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 1332/97
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 13

```
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-196-281-13

Query Match      67.0%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 7e-110;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDGRNDSSMHEAFRGKITALLYALGRTN 60
DB 29 NDVARGIVKADVAQNNGYLGQGVAVADTGLDGRNDSSMHEAFRGKITALLYALGRTN 88
QY 61 NANDPNHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPDLNLTLSQAWNA 120
DB 89 NANDPNHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPDLNLTLSQAWNA 148
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNDMTLVFAAGNPGNSGTISAPGTAKNAIT 180
DB 149 GARIHTNSGAPVNGAYTANSRQVDEYVRNDMTLVFAAGNPGNSGTISAPGTAKNAIT 208
QY 181 VGATENYRPSFGSLADPNHIAQFSRGATRDGRIPKDPVAPGTFTILSARSLAPDSFW 240
DB 209 VGATENYRPSFGSLADPNHIAQFSRGATRDGRIPKDPVAPGTFTILSARSLAPDSFW 268
QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFVKNRGTVPKPSLKAALIAAGATDVG- 300
DB 269 ANHDSKYAYMGTSMATPIVAGNVAQLREHFVKNRGTVPKPSLKAALIAAGATDVG- 328
QY 301 SGDOGWRVTLDKSLNV 317
DB 329 NGNQGWGRVTLDKSLNV 345

RESULT 14
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: TAKAKURA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-818B-1

Query Match      19.6%; Score 443; DB 3; Length 659;
Best Local Similarity 31.8%; Pred. No. 2.8e-26;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;

QY 8 VKADVAQNNGYLGQGVAVADTGLDGRNDSSMHEAFRGKITALLY-ALGRTNANDPN 66
DB 145 IGADTVNNSLGYDGSVVAIVDTGIDAN-----HPDLKGVIGWYDAVNGRSTPYDDQ 198
QY 67 GHGTHVAGSVLG-----NALNKGMAPOANLVFQSIM--DSSGGLGGLPDLNLTLSQAWNA 120
DB 199 GHGTHVAGSVLG-----NALNKGMAPOANLVFQSIM--DSSGGLGGLPDLNLTLSQAWNA 258
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNDMTLVFAAGNPGNSGTISAPGTAKNAIT 180
DB 259 GIRVINLSLSSQSSDGTDSLQAVNNADAGIIVCVVAGNSGPNVTYVGSPPAAASKVIT 318
QY 181 VGATENYRPSFGSLADPNHIAQFSRGATRDGRIPKDPVAPGTFTILSARSLAPDSFW 240
DB 319 VGA-----VDSNDNIASFSSRGPTADGRKLPDEVVAPGVDDIAPRAS---GTSMG 364
QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFVKNRGTVPKPSLKAALIAAGATDVG- 296
DB 365 TPINDYTKASGTSMATPHVSGVALILQ---AHPSTWPKVKTALITETADIVAPKEIAD 421
QY 297 LGYPSGQGWGRVTLDKSLNVAYVNEATATLQO-----KATYSFQAQAGKPLKISLWVT 351
DB 422 IAY-----GAGRVNVYKA--IKYDDYAKLTFTGSVADKGSATHTFDVSGATFTVATLYWD 474
QY 352 DAPGTTASVTLVNDLDELITAPNGQYKVGNDPSYPYDNNWDRNNVENVFINAPQSGTY 411
DB 475 -----TGSSDILLYLDPN-----GNEVDYSYATY---GFEKVGYNPTAGTW 515
QY 412 TIEVQAVNVPSGPQRFSLAIV 432
DB 516 TVKVVSY---KGAANYQVDVV 533

RESULT 15
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMJOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-09-445-472-12
```

Query Match 19.6%; Score 443; DB 4; Length 659;
Best Local Similarity 31.1%; Fred.No. 2.8e-26;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;

QY	8	VKADVAQNNYGLYGQGVAVADTGLDTRNDSSNHEAFRGKITALLY-ALGRTNANDPN	66
Db	145	IGADTVMSNLGYDGGGVVAIVDTGIDAN-----HPDLKKGKVIWYDAVNGRSTPYDDQ	198
QY	67	GHGTHVAGSVLG-----NALNKGMAPOANLVFQSIM--DSSGGLGGLPSNLNTLFSQAWNA	120
Db	199	GHGTHVAGIVAGTGSVNSQYIGVAFGAKLVGVKLGADGSGSVSTIIAGVDVVVQNKDKY	258
QY	121	GARIHTNSMGAPVNGAYTANSRQVDEYVVRNNDMTVLPFAAGNEGPNSGTISAPGTAKNAIT	180
Db	259	GIRVINLSLGSSQSSDGTDSLSCAVNNAWDAGIVVCVAAGNSGPNITYVGSPPAAASKVIT	318
QY	181	VGATENYRPSFGLADNPNHIAQFSRGATRDGRIPDVTAPGTFFILSARSLAPDSFW	240
Db	319	VGA-----VDSNDNIASFSSRGPTADGRLKPEVAPGVVDIIAPRAS---GTSMG	364
QY	241	ANVNSKYAVMGTSMATPIVAGNVAQLREHFKNRGITP---KPSLIKAALIAGATDVG-	296
Db	365	TFINDYYTKASGTSAATPHVSGVGLILO--AHPSWTFDKVKYKTALITADIAPKEIAD	421
QY	297	LGYPGDDQGWGRVTLDKSLNVAYVNEATALATGQ-----KATYFQAQAGKPLKLSLWT	351
Db	422	IAY-----GAGRNVVYKA--IKYDDYAKLFTFTGSVADKGSATHTFDVSGATFVTATLYWD	474
QY	352	DAPGSTTASVTLVNDLDLVTAPNGOKYVGNDFSPYVDNNWDGRNNVNFINAPOSGTY	411
Db	475	-----TGSSDIDLILYDPN-----GNEVDYSYAYY-----GFEKVGYYNPTAGTW	515
QY	412	TIEVQAYNVPSGPQRFSLAIV	432
Db	516	TVKWSY---KGAANYQVDVY	533

Search completed: March 31, 2004, 16:11:55
Job time : 14.9892 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:09:09 ; Search time 33.1412 Seconds
(without alignments)
3418.697 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVANNYGLY.....EVQAVNVPFGQRFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubaa/PTC_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubaa/PTCUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubaa/US09A_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2261	100.0	433	10 US-09-985-689A-5	Sequence 5, Appli
2	2251	99.6	433	10 US-09-985-689A-3	Sequence 3, Appli
3	2240	99.1	433	10 US-09-985-689A-4	Sequence 4, Appli
4	2044	90.4	433	10 US-09-985-689A-7	Sequence 7, Appli
5	2024	89.5	434	10 US-09-985-689A-6	Sequence 6, Appli
6	2010	88.9	434	10 US-09-985-689A-2	Sequence 2, Appli
7	1998.5	88.4	434	10 US-09-985-689A-1	Sequence 1, Appli
8	1998.5	88.4	434	15 US-10-385-662-2	Sequence 10, Appli
9	1514	67.0	345	14 US-10-336-324-10	Sequence 13, Appli
10	1514	67.0	345	14 US-10-403-105-13	Sequence 12, Appli
11	443	19.6	659	13 US-10-090-624-12	Sequence 1, Appli
12	422.5	18.7	412	13 US-10-090-624-1	Sequence 4, Appli
13	422.5	18.7	522	13 US-10-090-624-4	Sequence 16, Appli
14	422.5	18.7	654	13 US-10-090-624-16	Sequence 39, Appli
15	347.5	15.4	1079	14 US-10-112-488-39	

16	341.5	15.1	1237	14	US-10-314-657-4	Sequence 4, Appli
17	329	14.6	1139	14	US-10-156-761-10856	Sequence 10856, A
18	327	14.5	1208	14	US-10-156-761-13251	Sequence 13251, A
19	307	13.6	1398	13	US-10-090-624-6	Sequence 6, Appli
20	298.5	13.2	580	10	US-09-927-827-55	Sequence 55, Appli
21	295	13.0	519	15	US-10-084-846A-114	Sequence 114, Appli
22	295	13.0	19725	15	US-10-084-846A-4	Sequence 4, Appli
23	284.5	12.6	1101	14	US-10-156-761-12934	Sequence 12934, A
24	276.5	12.2	368	12	US-10-344-231-3	Sequence 3, Appli
25	276	12.2	271	10	US-09-813-408-2	Sequence 2, Appli
26	271	12.0	271	14	US-10-242-549-56	Sequence 56, Appli
27	270	11.9	271	14	US-10-242-549-54	Sequence 54, Appli
28	270	11.9	271	14	US-10-242-549-60	Sequence 60, Appli
29	269	11.9	289	8	US-08-322-678-10	Sequence 10, Appli
30	269	11.9	269	9	US-09-837-235-16	Sequence 16, Appli
31	269	11.9	269	9	US-09-060-854B-6	Sequence 6, Appli
32	269	11.9	269	9	US-09-975-139-1	Sequence 1, Appli
33	269	11.9	269	9	US-09-976-414-8	Sequence 8, Appli
34	269	11.9	269	10	US-09-736-116-49	Sequence 49, Appli
35	269	11.9	269	13	US-10-075-907-1	Sequence 1, Appli
36	269	11.9	269	13	US-10-075-895-1	Sequence 1, Appli
37	269	11.9	269	14	US-10-033-325-6	Sequence 6, Appli
38	269	11.9	269	14	US-10-209-812-3	Sequence 3, Appli
39	269	11.9	269	14	US-10-104-693-4	Sequence 4, Appli
40	269	11.9	269	14	US-10-228-572-6	Sequence 6, Appli
41	269	11.9	269	14	US-10-324-152-5	Sequence 5, Appli
42	269	11.9	269	14	US-10-242-549-5	Sequence 5, Appli
43	269	11.9	269	14	US-10-313-853-1	Sequence 1, Appli
44	269	11.9	269	14	US-10-336-324-4	Sequence 4, Appli
45	269	11.9	269	14	US-10-336-324-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-985-689A-5
; Sequence 5, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-385166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-5

Query Match 100.0%; Score 2261; DB 10; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVANNYGLYGGQGVAVADTGLDTGRNDSSMHEAFRGKITALYALGRTN 60

Db 1 NDVARGIVKADVANNYGLYGGQGVAVADTGLDTGRNDSSMHEAFRGKITALYALGRTN 60

QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIQNDSSGGIGGLPNSLNTLFSQAWNA 120

Db 61 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGPSNLNTLFSQAWNA 120
QY 121 GARIHTNSWGPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
Db 121 GARIHTNSWGPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
Db 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
QY 241 ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
Db 241 ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
QY 301 SGQDQGRVTLDKSLNVAYVNEATALATGKATYSFQAQAGKPKLSLWTDAPGSTTAS 360
Db 301 SGQDQGRVTLDKSLNVAYVNEATALATGKATYSFQAQAGKPKLSLWTDAPGSTTAS 360
QY 361 YTLVNDLVLVITAPNGQKYVGNDFSYPDYNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
Db 361 YTLVNDLVLVITAPNGQKYVGNDFSYPDYNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
QY 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

RESULT 2

US-09-985-689A-3
; Sequence 3, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR FILING DATE: 2000-11-22
; PRIOR FILING DATE: 2000-11-22
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-3

Query Match 99.6%; Score 2251; DB 10; Length 433;
Best Local Similarity 99.3%; Pred. No. 8.7e-191;
Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAPRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAPRGKITALYALGRTN 60
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGPSNLNTLFSQAWNA 120
Db 61 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGPSNLNTLFSQAWNA 120
QY 121 GARIHTNSWGPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
Db 121 GARIHTNSWGPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180

QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
Db 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
QY 241 ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
Db 241 ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
QY 301 SGQDQGRVTLDKSLNVAYVNEATALATGKATYSFQAQAGKPKLSLWTDAPGSTTAS 360
Db 301 SGQDQGRVTLDKSLNVAYVNEATALATGKATYSFQAQAGKPKLSLWTDAPGSTTAS 360
QY 361 YTLVNDLVLVITAPNGQKYVGNDFSYPDYNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
Db 361 YTLVNDLVLVITAPNGQKYVGNDFSYPDYNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
QY 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

RESULT 3

US-09-985-689A-4
; Sequence 4, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR FILING DATE: 2000-11-22
; PRIOR FILING DATE: 2000-11-22
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-4

Query Match 99.1%; Score 2240; DB 10; Length 433;
Best Local Similarity 98.8%; Pred. No. 8.2e-190;
Matches 428; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAPRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAPRGKITALYALGRTN 60
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGPSNLNTLFSQAWNA 120
Db 61 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGPSNLNTLFSQAWNA 120
QY 121 GARIHTNSWGPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
Db 121 GARIHTNSWGPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
Db 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
QY 241 ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
Db 241 ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300

```
QY 301 SDGQGRVTLDKSLNVAVNEATATATCQKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 NGDQGRVTLNKSUNVAVNEATATATCQKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
QY 361 YTLVNDLDLIVITAPNGQKTVGNDFSYYPYDNNWDGRNNVENVFAPQSGTYTIEVQAYNV 420
Db 361 YTLVNDLDLIVITAPNGQKTVGNDFSYYPYDNNWDGRNNVENVFAPQSGTYTIEVQAYNV 420
QY 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

RESULT 4
US-09-985-689A-7
; Sequence 7, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-7

Query Match 90.4%; Score 2044; DB 10; Length 433;
Best Local Similarity 89.1%; Pred. No. 1.9e-172; Indels 0; Gaps 0;
Matches 386; Conservative 24; Mismatches 23;

QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNAL-NKGWAPQANLVFQSTMDSSGGLGPSNLNTLFSQAWN 120
Db 61 NANDPNHGHTHVAGSVLGNATNKGWAPQANLVFQSTMDSSGGLGPSNLNTLFSQAYSA 120
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSGTISAPGTAKNAIT 180
Db 121 GARIHTNSGAPVNGAYTTDSRNVDDYVKNDMTVLFAGNEGPGSGTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGLADNPHIAQFSSRGATRGRIKPDVTPAGTTLISARSLAPDSSF 240
Db 181 VGATELRLPSFGSYADNINHVAQFSSRGPTRGRIKPDVMAFGTYILSARSLAPDSSF 240
QY 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALITAGATDVLGYP 300
Db 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFKVNGRVTPEKPSLLKXALJAGAADVLGFP 300
QY 301 SDGQGRVTLDKSLNVAVNEATATATCQKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 NGDQGRVTLDKSLNVAFVNETSPLSTSQKATYSFTAQAGKPLKISLVSDAPGSTTAS 360
QY 361 YTLVNDLDLIVITAPNGQKTVGNDFSYYPYDNNWDGRNNVENVFAPQSGTYTIEVQAYNV 420
Db 361 YTLVNDLDLIVITAPNGQKTVGNDFSYYPYDNNWDGRNNVENVFAPQSGTYTIEVQAYNV 420
```

```
Db 361 LTLVNDLDLIVITAPNGTKYVGNDFTA PYDNNWDGRNNVENVFAPQSGTYTIEVQAYNV 420
QY 421 PSGPQRFSLAIVH 433
Db 421 PVSFQTFSLAIVH 433

RESULT 5
US-09-985-689A-6
; Sequence 6, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-6

Query Match 89.5%; Score 2024.5; DB 10; Length 434;
Best Local Similarity 88.5%; Pred. No. 1e-170; Indels 1; Gaps 1;
Matches 384; Conservative 28; Mismatches 21;

QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 1 NDVARGIVKADVAQSSYGLYGGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNAL-NKGWAPQANLVFQSTMDSSGGLGPSNLNTLFSQAWN 119
Db 61 NANDPNHGHTHVAGSVLGNATNKGWAPQANLVFQSTMDSSGGLGPSNVSTLFSQAYS 120
QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSGTISAPGTAKNAI 179
Db 121 AGARIHTNSGAPVNGAYTTDSRNVDDYVKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
QY 180 TVGATENYRPSFGLADNPHIAQFSSRGATRGRIKPDVTPAGTTLISARSLAPDSSF 239
Db 181 TVGATELRLPSFGSYADNINHVAQFSSRGPTRGRIKPDVMAFGTYILSARSLAPDSSF 240
QY 240 WANYNSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALITAGATDVLGYP 299
Db 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLLKXALITAGATDVLGYP 300
QY 300 PSGDQGRVTLDKSLNVAVNEATATATCQKATYSFOAQAGKPLKISLVWTDAPGSTTAS 359
Db 301 PSGNQGRVTLDKSLNVAFVNETSPLSTNQKATYSFTAQSGKPLKISLVSDAPASTSA 360
QY 360 SYTLVNDLDLIVITAPNGQKTVGNDFSYYPYDNNWDGRNNVENVFAPQSGTYTIEVQAYN 419
Db 361 SVTLVNDLDLIVITAPNGTKYVGNDFTA PYDNNWDGRNNVENVFAPQSGTYTIEVQAYN 420
QY 420 VPSGPQRFSLAIVH 433
Db 421 VPQGPQAFSLAIVN 434
```

RESULT 6

US-09-985-689A-2
 ; Sequence 2, Application US/09985689A
 ; Publication No. US20030022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483USO
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 434
 ; TYPE: PRT
 ; ORGANISM: Bacillus sp.
 US-09-985-689A-2

Query Match 88.9%; Score 2010.5; DB 10; Length 434;
 Best Local Similarity 88.5%; Pred. No. 1.8e-169;
 Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;

QY	1	NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGTGRNDSSMHEAPRGKITALYALGRN	60
DB	1	NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGTGRNDSSMHEAPRGKITALYALGRN	60
QY	61	NANDPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFSQAWN	119
DB	61	NANDPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFSQAWN	120
QY	120	AGARIHNSWGPVNGAYTANSQVDEYVRNDMTVLFAAGNEGPNSTISAPGTAKNAI	179
DB	121	AGARIHNSWGAANGAYTTDSRVDDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI	180
QY	180	TVGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKDPVTAPGTFTLSARSSLAPDSSF	239
DB	181	TVGATENLRPSFGSADNINHVAFSSRGPTKGRIPKDPVMAFGTFTLSARSSLAPDSSF	240
QY	240	WANYNSKYAYMGTSMTPIVAGNVAQLREHFKNKGITPKPSLIKAALIAGATDVGLGY	299
DB	241	WANHDSKYAYMGTSMTPIVAGNVAQLREHFKNKGITPKPSLIKAALIAGAADVGLGY	300
QY	300	PSGQGWGRVTLDKSLNVAVNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTA	359
DB	301	PNGNQGWGRVTLDKSLNVAVNESSALSTSQKATYFTATAGKPLKISLVWSDAPASTTA	360
QY	360	SYTLVNDLDELVTAPNGQKTVGNDFSYPYNNWGRNNVENFINAPQSGTYYTIEVOAYN	419
DB	361	SVTLVNDLDELVTAPNGTRYVGNDFSPFNNWGRNNVENFINPQSGTYYTIEVOAYN	420
QY	420	VPSGPQRFSLAIVH 433	
DB	421	VPVGPQNFSLAIVN 434	

RESULT 7

US-09-985-689A-1
 ; Sequence 1, Application US/09985689A
 ; Publication No. US20030022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483USO
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 434
 ; TYPE: PRT
 ; ORGANISM: Bacillus sp.
 US-09-985-689A-1

Query Match 88.4%; Score 1998.5; DB 10; Length 434;
 Best Local Similarity 87.8%; Pred. No. 2.1e-168;
 Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY	1	NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGTGRNDSSMHEAPRGKITALYALGRN	60
DB	1	NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGTGRNDSSMHEAPRGKITALYALGRN	60
QY	61	NANDPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFSQAWN	119
DB	61	NANDPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFSQAWN	120
QY	120	AGARIHNSWGPVNGAYTANSQVDEYVRNDMTVLFAAGNEGPNSTISAPGTAKNAI	179
DB	121	AGARIHNSWGAANGAYTTDSRVDDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI	180
QY	180	TVGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKDPVTAPGTFTLSARSSLAPDSSF	239
DB	181	TVGATENLRPSFGSADNINHVAFSSRGPTKGRIPKDPVMAFGTFTLSARSSLAPDSSF	240
QY	240	WANYNSKYAYMGTSMTPIVAGNVAQLREHFKNKGITPKPSLIKAALIAGATDVGLGY	299
DB	241	WANHDSKYAYMGTSMTPIVAGNVAQLREHFKNKGITPKPSLIKAALIAGAADVGLGY	300
QY	300	PSGQGWGRVTLDKSLNVAVNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTA	359
DB	301	PNGNQGWGRVTLDKSLNVAVNESSALSTSQKATYSFTATAGKPLKISLVWSDAPASTTA	360
QY	360	SYTLVNDLDELVTAPNGQKTVGNDFSYPYNNWGRNNVENFINAPQSGTYYTIEVOAYN	419
DB	361	SVTLVNDLDELVTAPNGTRYVGNDFSPYNNWGRNNVENFINAPQSGTYYTIEVOAYN	420
QY	420	VPSGPQRFSLAIVH 433	
DB	421	VPVGPQNFSLAIVN 434	

RESULT 8

US-10-385-662-2
 ; Sequence 2, Application US/10385662
 ; Publication No. US20040002432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: SAITO, KAZUHIRO
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: IZAWA, YOSHIYUKI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; APPLICANT: KOBAYASHI, TOHRU
 ; APPLICANT: NOMURA, WAKAFUMI
 ; TITLE OF INVENTION: Alkaline protease

```
; FILE REFERENCE: 234938US0
; CURRENT APPLICATION NUMBER: US/10/385,662
; CURRENT FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: JP 2002-081428
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: JP 2002-165987
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: JP 2002-304230
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: JP 2002-304231
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp. KSM-KP43
US-10-385-662-2

Query Match      88.4%; Score 1998.5; DB 15; Length 434;
Best Local Similarity 87.8%; Pred. No. 2.1e-168;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYLYGGQVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGN-ALNKGMAPOANLVFQSIMDSSGGLGLPSNLTLFSQAWN 119
Db 61 NANDPNHGHTHVAGSVLGNSTNKGMAPOANLVFQSIMDSSGGLGLPSNLTLFSQAYS 120
QY 120 AGARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSTISAPGTAKNAI 179
Db 121 AGARIHTNSWGAAGVNGAYTTDSRNVDDYVRKNDMILFAAGNEGPNSTISAPGTAKNAI 180
QY 180 TVGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIPKDPVAPGTFILSARSSSLAPDSFF 239
Db 181 TVGATENLRPSFGSYADNINHAQFSSRGPTDGRIPKDPVAPGTFILSARSSSLAPDSFF 240
QY 240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLIKAALIAAGTADVGLGY 299
Db 241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLIKAALIAAGTADVGLGY 300
QY 300 PSGDQGWGRVTLDKSLNVAINEATATGQKATYSFOQAQPKLISLWTDAPGSTTA 359
Db 301 PNGNQGWGRVTLDKSLNVAINEATATGQKATYSFOQAQPKLISLWSDAPASTTA 360
QY 360 SVTLVNDLVLITAPNGQKVGNDPSYPVDNKGWGNVNVFNAPQSGTITIEVQAYN 419
Db 361 SVTLVNDLVLITAPNGTQVGNDFSTPNDNKGWGNVNVFNAPQSGTITIEVQAYN 420
QY 420 VPSGQPFSLAIHVH 433
Db 421 VPVGPQTFSIAIWN 434

RESULT 9
US-10-336-324-10
; Sequence 10, Application US/10336324
; Publication No. US20030176304A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349.204-US
; CURRENT APPLICATION NUMBER: US/10/336,324
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: US/09/512,251A
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-10-336-324-10

Query Match      67.0%; Score 1514; DB 14; Length 345;
Best Local Similarity 90.2%; Pred. No. 1.3e-125;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLYGGQVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNNGYLYGGQIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 88
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGLPSNLTLFSQAWN 120
Db 89 NANDPNHGHTHVAGSVLGNATNKGMAPOANLVFQSIMDSSGGLGLPANLQTLFSQAYS 148
QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSTISAPGTAKNAI 180

RESULT 10
US-10-403-105-13
; Sequence 13, Application US/10403105
; Publication No. US20030180933A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants And Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/10/403,105
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/196,281A
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1332/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-10-403-105-13

Query Match      67.0%; Score 1514; DB 14; Length 345;
Best Local Similarity 90.2%; Pred. No. 1.3e-125;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLYGGQVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNNGYLYGGQIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 88
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGLPSNLTLFSQAWN 120
Db 89 NANDPNHGHTHVAGSVLGNATNKGMAPOANLVFQSIMDSSGGLGLPANLQTLFSQAYS 148
QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSTISAPGTAKNAI 180
```

Db 149 GARIHTNSGAPVNGAYTTDSRNVDDYVAKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 208
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRIPKPDVTAGTFTILSARSSLAPDSSFW 240
Db 209 VGATENLRFSGSYADNINHVAQFSSRGTRDGRIPKPDVMAPTFTILSARSSLAPDSSFW 268
QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFHFKNGITPKPSLIKAALIAGATDVGLGYP 300
Db 269 ANHDSKYAYMGTSMATPIVAGNVAQLREHFHFKNGITPKPSLIKAALIAGAADVGLGFP 328
QY 301 SGDQGWGRVTLDKSLNV 317
Db 329 NGNQGWGRVTLDKSLNV 345

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 19.6%; Score 443; DB 13; Length 659;
Best Local Similarity 31.1%; Pred. No. 2,8e-30;
Matches 137; Conservative 66; Mismatches 170; Indels 69; Gaps 16;

QY 8 VKADVAQNNGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALLYALGRTNNAN-----DNGHG 69
Db 145 IGADTVNLSLGYDGSVVAIVDTGIDAN-----HPDLKGVIGWYDAVNGRSTPYDDQ 198

QY 67 GHGTHVAGSVLG-----NALNKGMAPOANLVFOSIM--DSSGGLGGLPSNLTLFQAANA 120
Db 199 GHGTHVAGSVLG-----NALNKGMAPOANLVFOSIM--DSSGGLGGLPSNLTLFQAANA 120

QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db 259 GIRVINLSLSSQSSDGTSLSAQVANNADAGIVVCAAGNPGNTYTVGSPAAASKVIT 318

QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRIPKPDVTAGTFTILSARSSLAPDSSFW 240
Db 319 VGA-----VDSNDNIASFSSRGPTADGRLKPEVAPGVPIIAPRAS---GTSMG 364

QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFHFKNGITP---KPSLIKAALIAGATDVG- 296
Db 365 TPINDYTKASGTSMATPHVSGVALILQ---AHPSWTPDKVKTALLETADIVAPKEIAD 421

QY 297 LGVPSDQGWGRVTLDKSLNVAAYNEATALATGQ-----KATYSFOAQAGKPLKISLVWT 351
Db 422 IAY-----GAGRVNYVKA--IKYDDYAKLTFTGSVADKGSATHETDVSATFTATLYWD 474

QY 352 DAPGSTTASVTLVNDLVLITAPNGQKYVGNDFSYPYDNNWMDGRNENVENFINAPQSGTYI 411
Db 475 -----TGSSDIDLVLIDPN-----GNEVDISYTAAY---GFEKGVGYNPTDGTW 515

QY 412 TIEVOANYVPSPQRFSLAIV 432
Db 516 TVKVVSY---KGAANYQVDVW 533

RESULT 12
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.7%; Score 422.5; DB 13; Length 412;
Best Local Similarity 31.4%; Pred. No. 9.1e-29;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALLYALGRTNNAN-----DNGHG 69
Db 20 NLGYDGSGITIGIDTGID-----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHG 68

QY 70 THVAGSVLG-----NALNKGMAPOANLVFOSIM--DSSGGLGGLPSNLTLFQAANA 122
Db 69 THVASIAAGTGAASNGKYKMAFGAKLAGIKVLGADGSGSISTIIKGVENAVDNCKYGI 128

QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAITVG 182
Db 129 KVINLSLSSQSSDGTSLSAQVAAWADAGLVVVAAGNSGPNKYTIGSPAAASKVITVG 188

QY 183 ATENYRPSFGSLADPNPHIAQFSSRGATRDGRIPKPDVTAGTFTILSARSSLAPDSSFWAN 242
Db 189 AVDKY-----DVITSFSSRGPTADGRLKPEVAPGVPIIAPRAS---GTSMGQP 234

QY 243 YNSKYAYMGTSMATPIVAGNVAQLREHFHFKNGITP---KPSLIKAALIAGATDVG-LG 298
Db 235 INDYTTAAPTGSMTATPHVAGIALLLQ---AHPSWTPDKVKTALLETADIVKPEIADIA 291

QY 299 YPSDQGWGRVTLDKSLNVAAYNEATALATGKA-----TYSFOAQAGKPLKISLVWTD 353
Db 292 Y-----GAGRVNYAKAIN--YDNYAKLVFTGYVANGSQTHQFVIGASGFVATLWDNA 344

QY 354 PGSTTASVTLVNDLVLITAPNGQKYVGNDFSYPYDNNWMDGRNENVENFINAPQSGTYI 413
Db 345 N-----SDDLVLVYDPNGQV---DYSY-----TAYYGFKEKGVGYNPTDGTWTI 385

QY 414 EVQANYVPSPQRFSLAIV 432
Db 386 KWSY---SGSANYQVDVW 401

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:

; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 ; FILE REFERENCE: TAKAKURA=6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/445,472
 ; PRIOR FILING DATE: 1999-12-06
 ; PRIOR APPLICATION NUMBER: 151969/1997
 ; PRIOR FILING DATE: 1997-06-10
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 522
 ; TYPE: PRT
 ; ORGANISM: Pyrococcus furiosus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (428)..(428)
 ; OTHER INFORMATION: Xaa at position 428 is Gly or Val.
 US-10-090-624-4

Query Match 18.7%; Score 422.5; DB 13; Length 522;
 Best Local Similarity 31.4%; Pred. No. 1.3e-28;
 Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;
 QY 16 NYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTNAN-----DPNGHG 69
 Db 20 NLGYDGGITIGIDTIGID-----ASHPLQGV-----IGWDFVNGRSYPYDDHGHG 68
 QY 70 THVAGSVLG-----NALNKGMAQANLVFQSIM--DSSGGLGGLPSNLTLFSQAWNAGA 122
 Db 69 THVASIAAGTGAASNGKYKGMAPCAKLAGIKVLGADGSGSISTIIKGVWAVDNKDKYGI 128
 QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSGTISAPGTAKNAITVG 182
 Db 129 KVINLSLGSSQSDGTDALSAQVNAADAGLVVVAAGSGPNKYITIGSPAAASKVITVG 188
 QY 183 ATENYRPSFGLADNPNHIAQFSRGATRDGRKPDVTPAGTFFILSARSLAPDSSFWAN 242
 Db 189 AVDKY-----DVIITFSRSGPTADGRLEKPEVAVPAGNWIIAARAS---GTSMGOP 234
 QY 243 YNSKYAYMGTSMATPTVAGNVAQLREHFINKRGITP---KPSLIKAALIAGATDVG-LG 298
 Db 235 INDYTTAAGTSMATPHVAGIAALLQ---AHPSTWPKVKKTALIEADIVKPEDEIADIA 291
 QY 299 YPSGDQGWRTLDKSLNVAYVNEATALATGOKA-----TYSFQAQAGKPLKISLVWTD 353
 Db 292 Y-----GAGRVNAYKAIN--YDNYAKLVFTGYVANKGSQTHQFVISCASFVTATLYWDNA 344
 QY 354 PGSTTASYTLVNDLIVITAPNGQKYVGNDFSPYDNNWGRNNVENVFINAQSGTYTI 413
 Db 345 N-----SDLDLYDPNGNQV---DYSY-----TAYYGFKEKGVYNTDGTGTWTI 385
 QY 414 EQVAYNVPSGPFPSLAIV 432
 Db 386 KWSY---SGSANYQVDVW 401

RESULT 14
 US-10-090-624-16
 ; Sequence 16, Application US/10090624
 ; Publication No. US20020132335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

; FILE REFERENCE: TAKAKURA=6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/445,472
 ; PRIOR FILING DATE: 1999-12-06
 ; PRIOR APPLICATION NUMBER: 151969/1997
 ; PRIOR FILING DATE: 1997-06-10
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 16
 ; LENGTH: 654
 ; TYPE: PRT
 ; ORGANISM: Pyrococcus furiosus
 US-10-090-624-16
 Query Match 18.7%; Score 422.5; DB 13; Length 654;
 Best Local Similarity 31.4%; Pred. No. 1.8e-28;
 Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;
 QY 16 NYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTNAN-----DPNGHG 69
 Db 152 NLGYDGGITIGIDTIGID-----ASHPLQGV-----IGWDFVNGRSYPYDDHGHG 200
 QY 70 THVAGSVLG-----NALNKGMAQANLVFQSIM--DSSGGLGGLPSNLTLFSQAWNAGA 122
 Db 201 THVASIAAGTGAASNGKYKGMAPCAKLAGIKVLGADGSGSISTIIKGVWAVDNKDKYGI 260
 QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSGTISAPGTAKNAITVG 182
 Db 261 KVINLSLGSSQSDGTDALSAQVNAADAGLVVVAAGSGPNKYITIGSPAAASKVITVG 320
 QY 183 ATENYRPSFGLADNPNHIAQFSRGATRDGRKPDVTPAGTFFILSARSLAPDSSFWAN 242
 Db 321 AVDKY-----DVIITFSRSGPTADGRLEKPEVAVPAGNWIIAARAS---GTSMGOP 365
 QY 243 YNSKYAYMGTSMATPTVAGNVAQLREHFINKRGITP---KPSLIKAALIAGATDVG-LG 298
 Db 367 INDYTTAAGTSMATPHVAGIAALLQ---AHPSTWPKVKKTALIEADIVKPEDEIADIA 423
 QY 299 YPSGDQGWRTLDKSLNVAYVNEATALATGOKA-----TYSFQAQAGKPLKISLVWTD 353
 Db 424 Y-----GAGRVNAYKAIN--YDNYAKLVFTGYVANKGSQTHQFVISCASFVTATLYWDNA 476
 QY 354 PGSTTASYTLVNDLIVITAPNGQKYVGNDFSPYDNNWGRNNVENVFINAQSGTYTI 413
 Db 477 N-----SDLDLYDPNGNQV---DYSY-----TAYYGFKEKGVYNTDGTGTWTI 517
 QY 414 EQVAYNVPSGPFPSLAIV 432
 Db 518 KWSY---SGSANYQVDVW 533
 RESULT 15
 US-10-112-488-39
 ; Sequence 39, Application US/10112488
 ; Publication No. US20030082746A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIKUCHI, Yoshiaki
 ; APPLICANT: DATE, Masayo
 ; APPLICANT: UMEZAWA, Yukiko
 ; APPLICANT: YOKOYAMA, Keiichi
 ; APPLICANT: MATSUI, Hiroshi
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING TRANSLUTAMINASE
 ; FILE REFERENCE: 219286USOCONT
 ; CURRENT APPLICATION NUMBER: US/10/112,488
 ; CURRENT FILING DATE: 2002-04-01
 ; PRIOR APPLICATION NUMBER: PCT/JPO00/06780
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: JP2000-280098
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: JP11-280098
 ; PRIOR FILING DATE: 1999-09-30
 ; NUMBER OF SEQ ID NOS: 70

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Streptomyces albogriseolus
US-10-112-488-39

Query Match      15.4%; Score 347.5; DB 14; Length 1079;
Best Local Similarity 29.7%; Pred. No. 1.7e-21;
Matches 130; Conservative 52; Mismatches 181; Indels 75; Gaps 16;

QY 18 GLYGOQVAVADTGLDGTGRNDSMEAFRGKITALYALGRNTNANDPNGHGHVAGSVL 77
Db 189 GYDGKGVKIAVLDTGVD-----ATHPLDKGVTAASKNFTSAPTGDVVGHGTHVASIAA 242
QY 78 GNALN-----KGMAPQANLVFQSIMDSGGGLGPSNLNTLFSQAWNA--GARIHTNSWG 130
Db 243 GTGAQSKGYKGVAPGAKILNGKVLDDA---GFGDDSGILAGMEWAAQAQADI VNMSLG 298
QY 131 A-----PVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAITVGA 183
Db 299 GMDTPETDPLEAA-----VDKLSAEKGILFAIAAGNEGPQS--IGSPGSADSALT VCA 349
QY 184 TENYRFSFGLADNPNIHQFSSRG--ATRDGRIKPDVTAPGTFFILSARSSILAPDSSFWAN 242
Db 350 -----VDDKDLADFSSTGPRLGDAVKPDLTAPGVDITAAASAKGNDIAKEVGE 398
QY 243 YNSKYAYMGTSMATPIVAGNVAQLREHFKNRGITP--KPSLIKAALIAGATDVGLGYP 300
Db 399 KPAGYMTISGTSWATPHVAGAAALLKQOH-----PENKYAELKGALTASTKDG--XYT 449
QY 301 SGDOGWGRVTLDXSLNVAVYNEATALATG-----QKATYSTQAQAGKPLKIS 347
Db 450 PFEQSGSRGVQVQDKAITQTIVIAEFVSLSGVQWPHADDXEVTKKLTyrNLGTEDVT LKLT 509
QY 348 LVMTDAPG-STTASYTLVNDLIVITAPNGCKYVGNDFSPYDNNMDGNNVNFVFINAP 406
Db 510 STATGPKKXAPAGFTLGLASTLTVPA-NGTASVDVTADTRLGGAVDGTYSAYVWATGAG 568
QY 407 QS-----GTYTIEVQAYNV 420
Db 569 QSVRTAAAREVESYNV 586
```

Search completed: March 31, 2004, 16:34:01
Job time : 34.1412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3246 Seconds
(without alignments)
3677.911 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQAYNVPSGQFQSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	501	22.2	1743	2	multidrug resistan
2	463.5	20.5	1905	2	multidrug resistan
3	344	15.2	444	2	intracellular alka
4	315.5	14.0	442	2	intracellular alka
5	312	13.8	806	2	microbial serine p
6	305	13.5	1398	2	pyrolysin (EC 3.4.
7	302.5	13.4	799	2	subtilisin-type pr
8	298.5	13.2	580	2	serine proteinase
9	285	12.6	1345	2	surface layer-asso
10	272.5	12.1	715	2	alkaline serine pr
11	269	11.9	380	2	high-alkaline seri
12	265	11.7	524	1	alkaline proteinas
13	263.5	11.7	420	1	subtilisin (EC 3.4
14	256.5	11.3	419	1	subtilisin (EC 3.4
15	256.5	11.3	1331	2	probable surface 1
16	255	11.3	401	2	serine proteinase
17	255	11.3	757	2	subtilisin-type pr
18	251.5	11.1	627	2	serine proteinase,
19	250	11.1	488	2	proteinase (import
20	250	11.1	894	2	cell wall-associat
21	244.5	10.8	1167	1	streptococcal Csa
22	243.5	10.8	378	2	high-alkaline seri
23	243	10.7	382	2	subtilisin (EC 3.4
24	242	10.7	379	1	subtilisin (EC 3.4
25	241.5	10.7	321	1	alkaline proteinas
26	241	10.7	513	1	aqualysin (EC 3.4.
27	238	10.5	384	2	alkaline proteinas
28	234.5	10.4	402	1	alkaline proteinas
29	232.5	10.3	319	2	microbial serine p

30	229.5	10.2	519	2	S71451	halolysin R4 (EC 3
31	228	10.1	1052	2	T17093	intraluminal subti
32	227.5	10.1	381	1	SUBSS	subtilisin (EC 3.4
33	227.5	10.1	381	2	JQ1487	subtilisin (EC 3.4
34	227.5	10.1	381	2	JH0778	subtilisin (EC 3.4
35	227.5	10.1	1374	2	D72593	hypothetical prote
36	226.5	10.0	381	1	SUBSI	subtilisin (EC 3.4
37	226.5	10.0	905	2	F82734	serine proteinase
38	226	10.0	613	2	S75976	hypothetical prote
39	225.5	10.0	1433	1	A36734	bacillopeptidase F
40	225	10.0	382	1	SUBSN	subtilisin (EC 3.4
41	224.5	9.9	272	2	A23624	subtilisin (EC 3.4
42	224	9.9	274	1	SUBSD	subtilisin (EC 3.4
43	224	9.9	601	2	JC4576	serine proteinase
44	223.5	9.9	279	1	SUMVTV	thermitase (EC 3.4
45	223	9.9	275	2	JC1085	subtilisin (EC 3.4

ALIGNMENTS

RESULT 1

T18279

multidrug resistance transport protein - slime mold (Dictyostelium discoideum)
C/Species: Dictyostelium discoideum

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T18279

R/Shaulsky, G.; Loomis, W.F.

submitted to the EMBL Data Library, June 1996

A/Reference number: Z18855

A/Accession: T18279

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1743 <SHA>

A/Cross-references: EMBL:U60086; NID:G1399914; PID:G1399915; PIDN:AAB03331.1

C/Genetics:

A/Gene: tagC

Query Match 22.2%; Score 501; DB 2; Length 1743;

Best local Similarity 27.7%; Pred. No. 2.5e-24;

Matches 163; Conservative 77; Mismatches 158; Indels 190; Gaps 24;

Qy	19	LYGGOVAVADTGLDTR	---NDS	-----SMHEAFRGKITALYALGRNTNANDPNGH	68
Db	314	LRKGQLLSIADTGLGSHCFSDSKYPIPLNSVNLNR	KVVTYITTTSDSDSKVDGH	372	
Qy	69	GTHVAGSVLG	---NALN	---KGMAPQANLVFQSIMDSSGGLGGLPSNLNLFQSAWNA	120
Db	373	GTHICGSAAGTPESSSVNISFSGLATDAKIAFFDLASGSSSLTP	PSDLKQLYQPLYDA	431	
Qy	121	GARIHNSWGA	---PVNGAYTANSRQVDEYV	RNNDMTVLPAAGNEGPNSTGIS	173
Db	432	GARVHCDSWGSVSGVTGTSYSSDTASIDFLTFHDFILRAAGN	---NEQYLSLLTQS	498	
Qy	174	TAKNAITVGGATE	-----NYRPS	-----FQS	193
Db	489	TAKNVIIVGAHQITHENYLTDGPNYINQYQSVVDINQELICDFDSRYCNVTTAQCLESNA	548		
Qy	194	-----	-----LADNPNHIAQFSSRGATRDGRKPDVTPGTILSA	229	
Db	549	TTGLASCCPTLLRKSRVIDAANTOPLYNENNICFSKSGKPTHGRMKPALVAPEVITSA	608		
Qy	230	RSSILA	-----PDSSFWANYNSKYAYMGTSMTPTVAGNVAQLREH	-----FIK	273
Db	609	RSNGANTTDCGGDSL	PNTNALLA	ISGTSMTATSPAAAATTLRQYLVVDVGYPTGSIVE	666
Qy	274	NRGTPKPSIIKAALIAGA	-----TDVGLGVPSGD	-----QMGVRLTD	312
Db	667	SNKLOPTGSLIKALMINNAQLNGTLITSSITYPNSQVFNENFAGASVQVQWGAIRMS	726		
Qy	313	KSLNVAVYNEAT	-----ALATGQKATYSF	-----	336
Db	727	NLHVANNNNNNKNTSDGITKFDGIGGLDLRLVKNQWKEBSLSLGTQNTSYCTYKPS	786		

A;Molecule type: DNA
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Residues: 1-442 <KUN>
A;Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PIDN:CAB13610.1; PID:el183385;
A;Experimental source: strain 168
C;Genetics:
A;Gene: aprX
C;Superfamily: subtilisin homology
F:146-398/Domain: subtilisin homology <SBT>

Query Match 14.0%; Score 315.5; DB 2; Length 442;
Best Local Similarity 28.6%; Pred. No. 4.2e-13;
Matches 103; Conservative 44; Mismatches 102; Indels 111; Gaps 13;
QY 11 DVAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALVALGRTNAN-----D 64
DB 136 EYVRNGQTLTGKGVAVVDTGI-----YHPDLEGR-----IGFADMVNQKTEPYD 183
QY 65 PNGHGHVAGSVLGNALN-----KGMAPQANLVFQSIMDSSG----- 101
DB 184 DNGHGHCHAGDVASSGSSGQYRGPAENLIGVKVLNKGSGTLADIISGVWCIOYN 243
QY 102 -----GLGG-----LPSNLTFLSQAWNAGARIHTNSWGAPVNGAYTAN 140
DB 244 EDNPPEDIDIMSLGGDALRDYDHEQEDPLVRAVEEWSAG----- 284
QY 141 SRQVDEYVRNDMTVLFAAGNPGNSGTISAPGTAKNALITVGATENYRPSFGSLADNPNH 200
DB 285 -----IVVCVAAGNSGPDQTIASPEVSEKVIITVGALDNN-----NTASSDDDT 328
QY 201 IAOFSRGATRDGRKPDYAPGTFTILSARSLAPDSF-----WANYNSKIYVNGGTSM 255
DB 329 VASFSRGTPTVYKKEPDIAPGVNIISLRS---PNSYIDKLQKSRVGSQYFTMSGTSM 385
QY 256 APIVAGNVQAUREHIKRGITPKSLIKAAIAGATDVLGYPSGQCGRGVTLDKSL 315
DB 386 ATPICAGIAALILQ---QNPDLTPDE--VK-ELLXNGTDKWKDEDNIFYGAGVNAENSV 439

RESULT 5

A41341
Microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 20-Jun-2000
C;Accession: A41341; B41341; S39700; D69730
J;Sloma, A.; Ruffo Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.
J. Bacteriol. 173, 6889-6895, 1991
A;Title: Cloning and characterization of the gene for an additional extracellular serine
A;Reference number: A41341; MUID:92041574; PMID:1938892
A;Accession: A41341
A;Molecule type: DNA
A;Residues: 1-806 <SLO>
A;Cross-references: GB:M76590; NID:g143819; PIDN:AAA22881.1; PID:g143820
A;Accession: B41341
A;Molecule type: protein
A;Residues: 161-195 <SL2>
R;Glaser, P.; Kunst, F.; Arnaut, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
A.; Rapoport, G.; Danchin, A.
Mol. Microbiol. 10, 371-384, 1993
A;Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
A;Reference number: S39655; MUID:95020537; PMID:7934828
A;Accession: S39700
A;Molecule type: DNA
A;Residues: 1-806 <GLA>
A;Cross-references: EMBL:X73124; NID:g413923; PIDN:CAA51601.1; PID:g580871
R;Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, C.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Kosono, S.; Hullo, M.F.
Koester, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
J.

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, R.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9884377
A;Accession: D69730
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-806 <KUN>
A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15835.1; PID:g2636344
A;Experimental source: strain 168
C;Comment: The amino terminal sequence of the mature protein and a molecular weight of
C;Genetics:
A;Gene: vpr
A;Start codon: TTG
C;Superfamily: microbial serine proteinase vpr; subtilisin homology
C;Keywords: hydrolase; serine proteinase
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-160/Domain: propeptide #status predicted <PRO>
F:180-548/Domain: subtilisin homology #status atypical <SBT>

Query Match 13.8%; Score 312; DB 2; Length 806;
Best Local Similarity 23.4%; Pred. No. 1.6e-12;
Matches 129; Conservative 56; Mismatches 156; Indels 210; Gaps 17;

QY 18 GLYGQGVAVADTGLDTCR-----NDSSMHEAFRGKITALVALGRTN 60
DB 177 GYTKGKIKVAITDGYEYVHPDLKKNFGQYGVDFVNDYDPKPTFG----- 224
QY 61 NAMDPNP-----HGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGLPSNLTLFSQ 116
DB 225 ---DPRCEATDHGTHVAGTVAANGTIKGVAPDATTLLAYRVLGPGG--SGTTENVIAGVER 279
QY 117 ANAGARIHTNSWGAPVNGAYTANSQVDEYVRNDMTVLFAAGNPGNSGTISAPGTAK 176
DB 280 AVQDQADWNLSLGNLNNPDWATSTALD--WAMSEGVAVTNSGNSGPNQWTVGSPGTSR 338
QY 177 NAITVGATE---NYRPSFGSL----- 194
DB 339 EATSVGATQLPLNEYAVTFGSSYSAKVMGYNKEDDVKALNNKEVELVEAGIGBAKDFEGK 398
QY 195 -----ADNP----- 198
DB 399 DLTKGVAWVRKGSIAFYDKADNAKAKAGAIQVYNNLSGSEIANVPGMSVFTKLSDLG 458
QY 199 -----NHIAQFSRKGATRD--GRIKPDVTAFTPTILSARSS 232
DB 459 EKLVSALKAGETKTTFKLTYSKALGEQVADFSSRGPVMDTMMIKPDISAFGVNIVTIPT 518
QY 233 LAPDSSFWANYNSKIYVNGGTSMATPIVAGNVQALREHFTKNRGITPKPSL--IKAAALIA 290
DB 519 HDPDHPY--GYGSKQ---GTSMAHPHAGAVAVIKQ-----AKPKWSVEQIKAAIMN 565
QY 291 GATDV---GLGYPSGQCGRGVTLDKSLNVAVYNEATATLGQKATY-SFQAQAGKPLK 345
DB 566 TAVTLKDSGDGVYPHNAQGAQ-----SARIMAKADSLVSPGSYSGYFTLKENGNETK 619
QY 346 ISLVMTDAPGSTTASYTLVNDLVLITAPNGQKYVGNDFSYPDYNNWGDGNNVENVFINA 405
DB 620 NETFTIENQSSIRKSYTL-----EYSFNGSGISTSGTSRVRVIPA 658
QY 406 PQSGTYTIEVQ 416
DB 659 HQTGKATAKVK 669

RESULT 6

T28159
Pyrolysins (EC 3.4.-.-) - Pyrococcus furiosus
C;Species: Pyrococcus furiosus
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000
C;Accession: T28159